Lecture Notes for *Philosophy & Science*

J. Dmitri Gallow

Draft of Spring, 2018

Contents

1 Course Intro

1. The philosophy of science subdivides into two broad categories:

   (a) *General* philosophy of science asks questions about (for instance) what the methodology of science is *in general*, whether/how it allows us to know things about the world, and what the scientific enterprise *in general* reveals or presupposes about the nature of reality.

   (b) *Applied* philosophy of science, on the other hand, asks questions about the methodology or the content of *particular* scientific theories. For instance, some questions in applied Philosophy of Science are:

      i. What is the best way to make sense of the theory of (non-relativistic) quantum mechanics? Is it one according to which the laws of nature are radically non-local—or is it one according to which the universe is constantly branching? Or is it something else altogether?

      ii. The Theory of Special Relativity tells us that whether two events are simultaneous depends upon how fast you are moving. This appears to fly in in the face of a traditional philosophical position known as *presentism*. Presentism claims that only the present is real—the past and the future are not real (at least, not any longer, or not yet). Is presentism really incompatible with special relativity? Or is there a way of making sense of the theory, and accomodating its predictions, which is consistent with presentism?

      iii. Most sciences make plentiful use of certain methods of statistical inference developed by the statisticians Fischer, Neyman, and Pearson. These methods are collectively known as *frequentist*. They are the methods taught in almost all undergraduate statistics courses. However, some philosophers and statisticians believe that these methods are flawed. There is a growing debate about which statistical procedures scientists ought to use. Who is right?

2. In this course, we will be focusing mainly on the *general* philosophy of science—though courses on applied philosophy of science are available through both the Philosophy department and the History and Philosophy of Science department here at Pittsburgh.

3. Questions in the general philosophy of science can be subdivided further, into *metaphysical* questions and *epistemological* questions.

(a) *Metaphysics* is the branch of philosophy which asks questions about which things exist, and *how* they exist. For instance,

   i. Do numbers exist? And, if so, what are they? Are they abstract entities which reside in a Platonic realm? Are they sets of concrete entities? Are they concepts? Are they just the nodes in a kind of abstract structure?
   ii. Do we have free will? Is free will compatible with our acts being causally determined?
   iii. Is a statue identical to the clay which makes it up? Or is the statue distinct from the clay which makes it up?

(b) *Epistemology* is the branch of philosophy which asks questions about how we may come to know things about the world, and whether our beliefs are *justified* or not. For instance,

   i. Can we know that we have hands? This seems like an easy thing to know—we can just look and see! But consider this: how can we know that we are not a brain in a vat, hooked up to a complicated computer simulation, being fed these experiences? If we can't rule this out, then it looks like perhaps we cannot rule out a possibility in which we don't have hands; so it begins to look like perhaps we cannot know that we have hands, after all.
   ii. In order to be know that, or be justified in believing that, we have hands, must we have evidence which rules out the possibility of error? Or can we know something (like, *e.g.*, that I skipped breakfast this morning) even when my evidence (like, *e.g.*, my memory) is consistent with that thing being false?
   iii. To what degree is it rational or reasonable to defer to expert opinion—like, *e.g.*, weather reporters or climatologists. What if you don't have any good independent reason to believe that they are reliable? Is it still rational to defer to them on matters about which you are ignorant?

4. One important task for the philosophy of science is answering what's known as the question of *demarcation*—that is, demarcating science from pseudo-science, or non-science.

   (a) Generally, scientists think that there is an important difference between the kind of intellectual activity that goes on in fields like Physics, Chemistry, Biology, and Cosmology, on the one hand, and the kind of intellectual activity that goes on in fields like Astrology, Phrenology, Intelligent Design (née Creation Science), and Marxism.
   (b) The difference, supposedly, is this: the former fields are *scientific*, while the later fields are *pseudo*-scientific. They pretend to be scientific, but in fact they are not. What does this difference amount to?

Part I

Induction

2 Naïve Inductivism, day 1

2.1 Inference

1. Suppose you have a collection of beliefs—call them $p_1, p_2, \ldots, p_N$. And, *on the basis of these beliefs*, you adopt a *new* belief—call it c.

   (a) In general, we'll put the propositions $p_1, p_2, \ldots, p_N$ *above* a horizontal line, and the proposition c *below* a horizontal line, to indicate that we are drawing an inference *from* $p_1, p_2, \ldots, p_N$, and *to* c.

   $$\frac{\begin{array}{c} p_1 \\ p_2 \\ \vdots \\ p_N \end{array}}{c}$$

   (b) In moving from $p_1, p_2, \ldots, p_N$ to c, you have drawn an *inference*.

   (c) We'll call $p_1, p_2, \ldots, p_N$ *premises*, and we'll call c a *conclusion*.

   (d) For instance,

   P1. Whoever committed the murder had the gate key

   P2. The butler had the gate key.

   C. The butler committed the murder.

2. Inferences can be divided into two kinds, which we will call *inductive* and *deductive*.[1]

   (a) In a *deductive* inference, the conclusion follows *necessarily* from the premises. More carefully, in a deductive inference, it is not possible for the premises to all be true at once, and yet for the conclusion to be simultaneously false.

   i. For instance, consider

   P1. The legs of an isosceles triangle have length l.

   C. The hypotenuse has length $\sqrt{2l^2}$.

[1] Beware: philosophical usage here is far from uniform. Some use 'ampliative' for the inferences we are calling 'inductive', and reserve 'inductive' for the inferences we will call 'enumerative inductions'.

ii. There is no way for the premise to be true while the conclusion is false. So this argument is *deductive*.

(b) On the other hand, in an *inductive* inference, the conclusion *doesn't* follow necessarily from the premises. Or, more carefully, in an inductive inference, it is possible for the premises to all be true at once, and yet for the conclusion to be simultaneously false.

i. For instance, consider

P1. All observed swans have been white.

C. All swans are white.

3. One particular kind of inductive inference is what we will call *enumerative induction*.

(a) For an example of an enumerative induction, suppose that we have observed N different ravens, and all of these ravens have been black. We might on that basis conclude that *all* ravens are black,

P1. The first raven is black.

P2. The second raven is black.

P3. The third raven is black.

⋮

PN. The Nth raven is black.

C. All ravens are black.

(b) In general, in an enumerative induction, we have premises stating that the first N Fs are Gs, and on this basis, we conclude that *all* Fs are Gs.

P1. The first F is G.

P2. The second F is G.

P3. The third F is G.

⋮

PN. The Nth F is G.

C. All Fs are G.

2.2 Naïve Inductivism

4. Francis Bacon's "new tool" (*novum organum*) of induction was a procedure for generating scientific knowledge inductively.

(a) His procedure consisted of 4 steps:

1. Observe and record facts.
2. Classify and analyze these facts.
3. Inductively derive generalizations from these facts.
4. Submit these generalizations to test.

(b) To understand these steps, let's walk through an example. Suppose that we observe the following facts:

- Clarence Oveur ate fish, drank coffee, & fell ill
- Roger Murdock ate fish, drank coffee, & fell ill
- Ted Striker ate chicken, drank soda, & didn't fall ill
- Elaine Dickinson ate chicken, drank coffee, & didn't fall ill
- Victor Basta ate fish, drank soda, & fell ill

(c) First, we should write these facts down (record them). Next, we should find some way of *classifying* these facts. For Bacon, this involved drawing up a table like the following:

Name	**Drink**	**Dinner**	**Health**
Oveur	Coffee	Fish	Sick
Murdock	Coffee	Fish	Sick
Striker	Soda	Chicken	Healthy
Dickinson	Coffee	Chicken	Healthy
Basta	Soda	Fish	Sick

(d) Once the facts have been classified, we should begin to *analyze* them. This involves looking for patterns. We notice that there is no firm correlation between coffee and health, nor soda and health. However, there is a correlation between fish and health. Everyone who ate fish got sick.

(e) So, next, form the analyzed data, we *inductively* derive generalizations. Using enumerative induction, we have

P1. Oveur ate fish and got sick.

P2. Murdock ate fish and got sick.

P3. Basta ate fish and got sick.

C. Everyone who eats the fish gets sick.

(f) Finally, we must submit our inductively derived generalization to a *test*. For instance, we could feed Striker the fish. If our hypothesis is true, then Striker will get sick. In Striker gets sick, our hypothesis has been *verified* (or *confirmed*). If Striker does not get sick, then our hypothesis has been *refuted* (or *disconfirmed*). In that case, we must start afesh.

5. For Bacon, it is very important that we complete stages 1 and 2 *without any preconceived hypotheses in mind*, and without any preconceived notions about which facts or classifications are important.

3 Naïve Inductivism, day 2

3.1 The Logic of Hypothesis Testing

1. In order to test a hypothesis, we must first find some *test implication* of the hypothesis.

    (a) For instance, suppose we have the following hypothesis: the reason water cannot be lifted above 10 meters with a vacuum is that water rises in a vacuum only because of the force exerted upon it by the atmospheric pressure. This hypothesis implies that, if we take our vacuum to a higher altitude, where the atmospheric pressure is lower, then we will only be able to lift water above 10 meters.

    (b) In general, in order to submit a hypothesis to test, we must discover something observation we would expect to make *if* the hypothesis were true. That is, in order to test a hypothesis H, we must discover some potential evidence E, such that we are assured of the following conditional:

    $$\text{If } H\text{, then } E$$

    (c) For instance, with respect to our hypothesis about the height of the water in a vacuum, in order to subject this hypothesis to tests, we must assure ourselves of this conditional:

    > If the hypothesis is true, then, at higher altitudes, we will be able to lift water higher than 10 feet with a vacuum pump

2. Once we have a *test implication*, we must go out and perform the test to see whether the hypothesis's implication in fact obtains.

    (a) For instance, we may bring a vacuum pump to the top of a high mountain and record how high water may be lifted there.

3. Suppose that the test implication is borne out—*e.g.*, suppose that a vacuum pump can raise water higher than 10 feet at the top of a high mountain. This is taken as a reason to accept the hypothesis. Formally, we will perform the following inference:

$$\begin{array}{l} \text{If } H\text{, then } E \\ E \\ \hline H \end{array}$$

4. On the other hand, suppose that the test implication is not borne out—*e.g*, suppose that the vaccum pump can *still* only raise water 10 feet at the top of the mountain, or suppose that it can only raise water 8 feet at the top of the mountain. This is taken as a reason to reject the hypothesis. Formally, we will perform the following inference:

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

5. We will be thinking quite a bit about the logical form of these inferences as the class progresses.

3.2 Hempel's Criticisms of Naïve Inductivism

6. Hempel begins with a story about the development of hand washing as a technique of disinfection.

   (a) Ignaz Semmelweis, a physician working in Vienna during the 1840's, realized that the rate of childbed fever was significantly higher in the 1st division than in the 2nd division.

   (b) To investigate why this was the case, he began by considering hypotheses. A common hypothesis at the time was that overcrowding causes childbed fever. However, upon investigation, it turns out that the 2nd division is more crowded than the 1st. So Semmelweis rejects this hypothesis.

   (c) He additionally considers the hypothesis that harsh treatment by the medical students is responsible for childbed fever (since the medical students are all trained in the 1st division). However, the rates of childbed fever remained constant after reducing the number of examinations. So Semmelweis rejects this hypothesis as well.

   (d) There are other hypotheses—like, *e.g.*, that the presence of a priest delivering last rights in the 1st division causes distress, which causes childbed fever. This hypothesis is rejected once Semmelweis has the priest take a different route, but the rates of childbed fever remain constant.

   (e) Finally, Semmeweis sees a colleague get cut with a scalpel while performing an autopsy. The colleague subsequently developed the same symptoms as childbed fever. Semmelweis notes that, in the 1st division, children were delivered by doctors who performed autopsies, while, in the second, they were delivered by midwives, who were not. This leads him to formulate the hypothesis that contamination with 'cadaveric matter' causes childbed fever. To test this hypothesis, Semmelweis has doctors cleanse their hands before deliveries. After this change, the rate of childbed fever in the 1st division dropped.

7. Recall the steps of Naïve Inductivism:

1. Observe and collect facts (*without a preconceived hypothesis in mind*).
2. Classify and analyze these facts (*without a preconceived hypothesis in mind*).
3. Inductively derive generalizations from these facts.
4. Submit these generalizations to test.

(a) Hempel objects to the first two steps of Naïve Inductivism. In particular, he objects to the parenthetical imperative to not let your observation and analysis be directed by any particular hypothesis.

(b) First, consider the injunction to collect facts without a preconceived hypothesis in mind.

i. In Semmelweis's case, the evidence he collected depended crucially upon the hypothesis he was considering. He considered how crowded the 1st and 2nd divisions were when he was entertaining the hypothesis that overcrowding causes childbed fever. Such information was irrelevant when he was considering the hypotheses that harsh treatment, the priest, or cadaveric matter causes childbed fever.

ii. Moreover, this feature of Semmelweis's investigation was, Hempel alleges, unavoidable. For there are an infinite number of facts. We must have some idea of where to start collecting facts, and some idea of which facts are relevant. And the natural place to look for guidance is the hypotheses we are considering.

(c) Next, consider the injunction to classify and analyze facts without a preconceived hypothesis in mind.

i. In Semmelweis's case, he could have classified women by age, fingernail length, eye color, *etc*. Instead, how he classified them depended crucially upon the hypothesis he was considering. When he was considering the hypothesis that harsh treatment by the medical students was causing childbed fever, he classified women by whether they were receiving examinations. When he was considering whether seeing the priest was responsible, he classified them by whether they had seen the priest or not.

ii. In general, there are infinitely many ways of classifying the facts. We must have some guidance as to which classifications might be relevant. And the natural place to look for this guidance is the hypotheses we are considering.

8. So, Hempel concludes: tentative hypotheses are required in order for scientific investigation to proceed. And these tenative hypotheses must be introduced *prior* to the collection or analysis of facts.

(a) Moreover, there can be no algorithmic method for selecting these initial hypotheses. Semmelweis needed the inspiration of his colleague's accident during the autopsy to have the correct hypothesis occur to him.

(b) Hempel concludes that science, just like mathematics, requires imagination and creativity.

9. Note, however, that, while this means that there is no algorithmic procedure for arriving at hypotheses, there could still be an algorithmic procedure for deciding whether or not to *accept* these hypotheses.

    (a) Hempel draws a distinction between the *context of discovery* and the *context of justification*.

    (b) The context of discovery is the stage of scientific investigation at which the scientist *discovers* a particular scientific hypothesis. For instance, Semmelweis discovered the hypothesis that cadaveric matter causes childbed fever upon observing his colleague's symptoms after cutting himself during the autopsy.

    (c) On the other hand, the context of *justification* is the stage of scientific investigation at which the scientist *justifies* believing in or accepting the hypothesis.

    (d) While Hempel does not believe that there is any method for telling you how to proceed in the context of discovery, there still may be, on his view, strict canons of inductive inference which tell you when a hypothesis has been *verified* (or *supported* or *confirmed*) by the available evidence. That is: there still may be a method to be followed in the context of justification.

4 The Problem of Induction, day 1

4.1 Some Prefatory Terminology

1. Let us divide our beliefs in two. On the one hand are those beliefs which are *a priori*, and on the other are those which are *a posteriori*.

   (a) A belief is *a priori* iff it can be known to be true *independent* of sense experience.

      i. For instance, my belief that 2+2=4 can be known to be true, independent of sense experience. Similarly, my belief that the sum of the squares of the legs of a right triangle is equal to the square of the hypotenuse can be known to be true, independent of sense experience.

      ii. What do we mean by 'independent of sense experience'? Think of it this way: in order to know the Pythagorean Theorem, you don't need to know that you're not in the Matrix. Even if you are plugged into the Matrix, the proof of the Pythagorean Theorem allows you to know it. Similarly, even if you are plugged into the Matrix, you still know that 2+2=4.

      iii. The important point is this: if a belief is *a priori*, then your justification for believing it does not rely upon your sense experience being *correct*. (Perhaps you have to have some sense experience at all in order to have the concept of a right triangle or of the numbers 2 and 4—but that sense experience doesn't have to be *accurate* in order for the belief that 2+2=4 to be justified.)

      iv. Note also that I don't say that a belief is *a priori* iff it *is* known independent of sense experience. I could come to know that 1234+4321=5555 by punching it into a calculator. Then, the basis of my belief would be that the calculator says that 1234+4321=5555. And, if I were to discover that I was plugged into the Matrix, then this wouldn't provide me with justification for believing that 1234+4321=5555. That doesn't make the belief not *a priori*. Even though the belief *isn't* known independent of sense experience, it still *can* be known independent of sense experience.

   (b) On the other hand, a belief is *a posteriori* iff it can only be known to be true through sense experience.

      i. For instance, my beliefs that Trump is president, that I am in Pittsburgh, and that it is sunny outside today are *a posteriori*. In order to come to know that Trump is president, that I am in Pittsburgh, and that it is sunny outside today, I must rely upon my sense experience.

ii. If I were to learn that I was plugged into the Matrix, then I would not be in a position to know that Trump is president, that I am in Pittsburgh, nor that it is sunny today.

2. Here's another way we may divide our beliefs:[1] on the one hand, there are the beliefs that are *necessary*. And, on the other hand, there are the beliefs which are *contingent*.

   (a) A belief is *necessary* iff there is no possible way for it to be false.

      i. For instance, my belief that there is not a square circle on display at the Smithsonian is *necessary*. There is no possible way for there to be a square circle, and so there is no possible way for my belief to be false.

      ii. Here, we do not mean only that there's no possible way—consistent with the laws of nature—for the belief to be false. We rather mean that there is no way *period* for the belief to be false.

   (b) A belief is *contingent* iff there is a possible way for it to be false, and a possible way for it to be true.

      i. For instance, my belief that the Earth rotates around the Sun approximately once every 365.25 days is *contingent*. It could be true, and it could be false.

4.2 Hume on Induction

3. Hume thinks that some of our beliefs are beliefs that we come by through noticing relations between our ideas. He calls these beliefs *Relations of Ideas*.

   (a) These are beliefs, like "All bachelors are unmarried" and "No woman is taller than herself", which we can know to be true just by consulting our ideas of "bachelor" and "taller than".

   (b) Included in this category will be all of mathematics. To know the Pythagorean Theorem, we need only consult our *idea* of a right triangle.

   (c) For Hume, these beliefs are discoverable by reason alone; and, for this reason, these beliefs are *a priori*.

   (d) Similarly, we cannot conceive of these beliefs being false. Since conceivability is our guide to possibility (it is possible iff it is conceivable), these beliefs could not possibly be false. So these beliefs are *necessary*, as well.

4. All beliefs which are not *Relations of Ideas*, Hume calls beliefs about *Matters of Fact*.

   (a) These are beliefs, like "Trump is president", "I am in Pittsburgh", and "The Earth travels around the Sun approximately once every 365.25 days". In order to come by these beliefs, we need to have some sense experience.

[1] Attentive students will note that this classification is not *exhaustive*—that is to say, we may have some beliefs which fall into neither category.

(b) Included in this category will be all of our beliefs about the natural world. In order to know that, *e.g.*, fire produces smoke, we need to observe fire. And in order to know that bread nourishes, *e.g.*, we need to eat some bread.

(c) These beliefs are not discoverable by reason alone; and, for this reason, these beliefs are *a posteriori*.

(d) We can easily conceive of these beliefs being false. We can conceive of Trump not winning the election; we can conceive of me not being in Pittsburgh, and we can conceive of the Earth taking longer to travel around the Sun. So these beliefs are *contingent*.

5. So, Hume believed that our two distinctions—between the *a priori* and the *a posteriori*, on the one hand, and the *necessary* and the *contingent*, on the other—lined up perfectly. He thought that a belief was *a priori* iff it is *necessary* and that a belief was *a posteriori* iff it was *contingent*. The former variety of beliefs are just the relations of ideas. The latter variety are just the matters of fact.

6. Even though some of our beliefs about matters of fact are based directly on sense experience (or memory), most of our beliefs about matters of fact are not.

   (a) For instance, my belief that Wesley Salmon wrote the article assigned for next class is not based directly upon sense experience. I didn't see him write it (how could I? He wrote it before I was born.).

   (b) Similarly, my belief that there will be a talk in the Philosophy department this coming Friday is not based directly upon sensne experience. I haven't seen the talk yet (how could I? It's in the future.).

   (c) I view smoke in the distance and believe that there is a fire there. I haven't seen the fire, so this belief is not based directly upon sense experience of the fire.

7. So, Hume enquires: 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*.

   (b) I believe that Wesley Salmon's writing the article would cause (somehow) his name to appear under the title. From this, and the fact that the name does appear there, I infer that Wesley Salmon must have written the article.

   (c) Similarly, I believe that inviting speakers and organizing talks causes those talks to take place. And I have seen this organization. So I infer that the talk will take place.

   (d) And I believe that fire causes smoke. So, when I see the smoke, I infer from the effect to the cause, and believe that there is fire causing the smoke.

8. So Hume has discovered an answer to his question: our beliefs about unobserved matters of fact are based on things we *have* observed, together with relations of cause and effect. Hume is not placated. He pushes on: what, then, is the basis of our beliefs about relations of cause and effect? Why do we believe that writing articles causes your name to appear beneath the title? Why do we believe that inviting speakers

and organizing talks causes those talks to take place? Why do we believe that fire causes smoke?

(a) Hume thinks that there are two possible ways for these beliefs to be known: either through reason, or through experience.

(b) But we can conceive of these beliefs being false. We can conceive of fire not causing smoke, or causing something else altogether. We can conceive of writing articles causing the name of your parents to appear beneath the title. We can conceive of organizing talks causing parties. If we can conceive of these beliefs being false, then they are contingent. But beliefs discoverable by reason alone must be *necessary*. Since these belies are not necessary, they cannot be discoverable by reason.

(c) So they must be discoverable by experience. Which experience justifies us in accepting these beliefs about cause and effect?

9. Hume's answer is that we justify our beliefs about relations of cause and effect through what we earlier in the semester called *enumerative induction*.

   (a) For instance, we have observed on several occasions that fire is followed by smoke. From this, we infer that fire is always followed by smoke.

   P1. The first fire was followed by smoke.

   P2. The second fire was followed by smoke.

   ⋮

   PN. The Nth fire was followed by smoke.

   C. Fire is always followed by smoke.

10. But, Hume asks, what justifies us in drawing *this* inference? It is surely conceivable that the first N fires be followed by smoke, yet the $N + 1$st fire fail to produce any smoke at all—or that the $N + 1$st smoke would fail to be produced by fire.

    (a) If it is conceivable that the premises be true, yet the conclusion false, then the inference cannot be determined by reason. For this means that the conclusion does not *necessarily* follow from the premises.

    (b) So, the inference cannot be justified on the basis of reason alone. Rather, it must be justified on the basis of some other beliefs we have.

11. So Hume enquires further: what other belief of ours could justify us in drawing the inference above?

    (a) His answer: we believe *not only* that the first N fires were followed by smoke. We additionally believe that nature operates in a uniform manner. We additionally believe what Hume calls the *principle of the uniformity of nature*.

    (b) And, if we supplement the reasoning above with this additional premise, P1. The first fire was followed by smoke.

P2. The second fire was followed by smoke.

⋮

PN. The Nth fire was followed by smoke.

UN. Nature operates uniformly.

C. Fire is always followed by smoke.

then Hume is happy to accept that reason alone could determine the inference that Fire is always followed by smoke.

(c) However, Hume is not so easily satisfied. From where did we derive *this* belief? It is either a belief about relations of ideas, or a belief about matters of fact. We can conceive of it being false, wherefore we must conclude that it is not necessary, and therefore, not a relation of ideas. So it must itself be a matter of fact.

12. So the belief that nature operates uniformly is itself a matter of fact. We have not directly observed nature operating uniformly in all of the unobserved cases. So what is our basis for *this* belief?

    (a) The most natural answer, it seems, is this: in all of the *observed* cases, nature has operated in a uniform manner. Thus far, when we go looking for regularities in nature, we have found them, and those discovered regularities have, for the most part, continued into the future.

    (b) So this is the basis of our belief that nature will operate uniformly, then: because it has always done so in the past.

    P1. Nature has always operated uniformly in the past.

    UN. Nature always operates uniformly

13. *But*—and this is Hume's first punchline—we can easily conceive of this premise being true while the conclusion is nevertheless false. We can easily conceive of nature operating uniformly up to the year 2020, and then, thereafter, operating in a haphazard, nonuniform manner.

    (a) There is no contradiction involved in the supposition that, on Jan. 1, 2020, massive bodies will begin *repelling* each other instead of *attracting* each other.

    (b) Nor is there any contradiction involved in the supposition that, on Jan 1, 2020, bread will stop nourishing and begin poisoning, or that it will keep nourishing blue eyed people and poison brown eyed people.

    (c) So, our past experience of nature operating uniformly does not provide a conclusive reason to suppose that nature will *continue* operating uniformly in the future.

14. So, Hume's first lesson is this: we cannot know by reasoning from what we've observed that our method of forming beliefs about the future will not lead us into radical error. We cannot know by reasoning from what we've observed that our scientific theories will continue to be true in the future.

5 Laws of Nature, day 1

1. Hume has argued that, for all we've observed, it could be that nature will stop operating uniformly tomorrow. The claim that nature operates uniformly was, for Hume, the claim that like causes will continue to produce like effects. However, nowadays, we are more likely to think of the claim that nature operates uniformly as the claim that nature operates according to certain *laws*.

2. Transposed into this key, Hume's claim is that, for all we've observed, it could be that the laws we *think* are universally obeyed could cease to hold tomorrow.

3. Let's take this opportunity to think a bit about what it is that might make something a law of nature.

5.1 Laws and Accidentally True Universal Generalizations

4. Some candidate laws of nature are:

   (a) Energy is always conserved

   (b) All bodies accelerate proportionally to the net forces acting upon them, and inversely to the total mass of the body ($F = ma$)

   (c) All electrons repel each other

   (d) Nothing may travel faster than the speed of light

5. The statements above are candidate laws of nature—but what is it for something to be a law of nature? What's the difference between the following two claims?

   (a) All electrons repel each other.

   (b) It is a law of nature that all electrons repel each other.

   What does 'it is a law that' add to these claims?

6. As a first pass answer to our question, we may note that it appears as though laws of nature are *exceptionless*. They are *universally true* generalizations.

(a) That is, the form that a law takes is this: *All Fs are Gs*, or *No Fs are Gs*. (These are universal generalizations.)

7. However, this first pass characterization won't do. For there seems to be a difference between the following two claims (both of which, we may suppose, are true universal generalizations):

   (a) All electrons repel each other.

   (b) All fruits in Smith's garden are apples.

   The first of these universal generalizations is a law. The second is merely an *accidentally true* universal generalizations. What makes the difference? In virtue of what are some universal generalizations *laws of nature*, and others mere *accidents*?

8. We might think that the problem with "All fruits in Smith's garden are apples" is that it, unlike "All electrons repel each other", makes reference to a specific person, Smith, and a specific place, his garden.

   (a) A new suggestion, then: in order to be a law of nature, a universally true generalization must not make reference to any specific person, place, or thing.

   (b) Let's call a property which does not make reference to any specific person, place, or thing a *qualitative* property.

   (c) Then, our new suggestion is this: a law of nature is just a universal generalization which only uses *qualitative* properties.

9. The general strategy we've adopted here is the following: we want to say that a law is a particular kind of universal generalization. So we are saying that

   $$\text{Law} \;=\; \text{Universal Generalization} \;+ X$$

   (a) And, the present proposal is that we let X be 'contains only qualitative properties'.

10. Unfortunately, the restriction to qualitative properties doesn't seem to help. For there are seemingly accidental generalizations which contain only qualitative predicates. For instance, "There are no gold spheres greater than 1 mile in diameter". This is a true universal generalization, but it is only accidentally true. There is no law which would prevent us from constructing a gold sphere more than 1 mile in diameter.

11. Contrast the following two claims:

   (a) There are no gold spheres greater than 1 mile in diameter.

   (b) There are no uranium spheres greater than 1 mile in diameter.

   While the first is accidentally true, the second is a law of nature. The critical mass of uranium makes it impossible, as a matter of law, for there to be uranium spheres that large. So, whatever the difference is between laws and merely accidentally true universal generalizations, it does not have to do with the properties being purely qualitative.

6 Laws of Nature, day 2

1. Last class, we considered the following question: what is it for a universal generalization (like "All Fs are Gs" or "No Fs are Gs") to be a law of nature? What, that is, is the difference between the following two claims?

   (a) All Fs are Gs

   (b) It is a law of nature that all Fs are Gs

2. Such an account should explain the difference between laws of nature and merely *accidentally true* universal generalizations. That is, such an account should tell us why (2a) is a law of nature, while (2b) is not.

   (a) There are no gold spheres larger than a mile in diameter.

   (b) There are no uranium spheres larger than a mile in diameter.

6.1 The "Best Systems" Account of Laws

3. Recall, last class we were considering accounts on which laws of nature are just a particular kind of universal generalization.

$$\text{Law} \;=\; \text{Universal Generalization} \;+ X$$

   (a) We considered allowing X to be "contains only *qualitative* properties (properties which do not make reference to any particular place, person, or time).

4. The first theory of laws of nature we're going to talk about today lets X be "is a consequence of the systematization of truths which strikes the best balance of simplicity and strength".

5. Suppose that it is your task to *systematize* the truths of our world—your goal is to write down a collection of truths, from which we may derive all kinds of interesting facts about our world. There are a variety of different systems you could write down.

   (a) Some of these systems are *simple*. They contain few and simple truths. Some of these systems are *complex*. They contain many and complicated truths.

(b) Some of these systems are *strong*. They tell us much about the world. Some of these systems are *weak*. They do not tell us much of anything about the world.

(c) Simplicity is a virtue of a systematization of truths. So too is strength. However, these virtues compete. As you make your system stronger, it will become more complicated. And as you make your system more simple, it will become weaker.

6. The proponents of the "Best Systems" account of laws of nature say that there is a system which strikes the *best balance* of simplicity and strength. And the laws of nature are those universal generalizations which are included in, or are derivable from, this *best system*.

   (a) On this account, to be a law is to be a particularly simple and informative universal generalization—or the consequence of such a simple, informative generalization.

   (b) The laws are what you could learn from reading the most concise and informative executive summary of the truths of our universe.

7. To think through some of the consequences of this account of laws, let's consider some examples.

   (a) On this account, it is impossible for there to be a Newtonian universe containing just a single particle which remains at rest for all time. Why? Well, if there were just that single particle, remaining at rest, then a system which contained $F = ma$ and the law of gravitation would be needlessly complex. Such a system could be greatly simplified by replacing the laws of Newtonian mechanics with the truth "the particle always remains at rest".

   (b) Consider a possible universe containing three kinds of particles: α particles, β particles, and γ particles. Whenever an α particle interacts with a β particle, they are both annihilated. Whenever a β particle interacts with a γ particle, they are both annihilated. However, it turns out that α and γ particles never interact.

      i. The "Best Systems" account will plausibly say that it is a law that α particles and β particles annihilate each other. This one, simple truth tells us much about the world. Similarly, it will plausibly say that it is a law that β particles and γ particles annihilate each other.

      ii. Intuitively, we've not yet said enough to know what the laws say about the interactions of α and γ particles. Intuitively, it might be a law that they annihilate each other, and it might be a law that they produce a β particle when they interact. Or maybe the laws say that, when they interact, they create a new kind of particle, a δ particle. Intuitively, all of these are genuine possibilities left open by the description above.

      iii. However, according to the "Best Systems" account of laws, these are not genuine possibilities. The laws are no more than executive summaries of the goings-on of a universe. If there are no interactions between α and γ particles, then there are no truths about those interactions to be summarized. Perhaps it will end up being a law that α and γ particles annihilate, since "Different kinds of particles annihilate each other" will be a simpler system than one which specifies that α and β particles annihilate and also β and γ particles annihilate. However, it couldn't possibly be a law that α and γ particles spawn some new particle.

8. An objection to the 'Best Systems' account:

    (a) Laws are generally taken to explain regularities. Why are all Fs Gs? Because it's a law that all Fs are Gs.

    (b) But, according to the 'Best Systems' account, regularities also explain the laws. Why is it a law that all Fs are Gs? Because all Fs are Gs (and this is a particularly simple and informative fact about the world).

    But then, it looks like the explanations we've provided are *circular*—laws explain regularities and regularities explain laws. If we think that there can't be circular explanations like this, then this is a problem for the 'Best Systems' account.

6.2 The Universals Account of Laws

9. There is another account of laws which is more realist about laws. According to this account, laws are something *over and above* the regularities we observe.

10. To introduce the account, a bit of metaphysical background:

    (a) Take two red apples. It seems that the apples have something in common—namely, their redness. This property of redness cannot be wholly present in either apple, since then, the two apples would not *share* the property of redness. So there must be a property of redness which is not wholly present in either apple, but which both apples *exemplify*. This property of redness must be something *over and above* the particular apples.

    (b) The *Platonist* accepts the foregoing reasoning. They think that, in addition to the various particulars of our world, there are additionally *universals*, which are over and above all of the particulars of our world (they live in 'Plato's heaven'). According to the Platonist, to say that the apple is red is to say that the *particular* apple *instantiates* the universal Redness. It is because both particular apples instantiate the very same universal that they have a property in common.

    (c) Not all philosophers are Platonists. The *nominalist* denies that there is any such thing as a universal Redness. According to the nominalist, there are only particulars.

11. Some philosophers (in particular, Armstrong, Dretske, and Tooley) contend that laws of nature are not about particulars at all. Rather, they are about *universals*.

    (a) On this account, to say that it is a law that all Fs are Gs is to say that there is a *necessitation* relation between the universal F-ness and the universal G-ness.

(b) When there is a necessitation relation like this between the universals F-ness and G-ness, then this will require that all Fs are Gs, though the converse is false—it could be that all Fs are Gs, even though F-ness doesn't necessitate G-ness. According to the Universals account, this is the difference between laws and merely accidentally true universal generalizations.

12. Let's think through how the Universals account deals with the examples we considered with the "Best Systems" account above:

    (a) On the universals account, it is possible for there to be a Newtonian universe containing just a single particle which remains at rest for all time. At that universe, there is a necessitation relation between the properties of being subject to such-and-such force, having such-and-such mass, and having such-and-such acceleration. Those laws govern in that universe, even though there are no forces there to govern.

    (b) Similarly, and for the very same reason, on the universals account, it is possible for there to be a universe with laws governing particle interactions, even when those particles never interact.

        i. Consider the possible universe containing three particles: α, β, and γ particles. it this universe, α and β particles interact, and β and γ particles interact, but α and γ particles never interact.

        ii. If there are the right necessitation relations amongst universals, then it may be a law that α and γ particles annihilate when they interact with each other, even though this never happens.

        iii. Similarly, if there are the right necessitation relations amongst universals, then it may be a law that α and γ particles spawn a δ particle if they interact, even though this never happens.

        iv. More generally, for any law you can imagine governing the interactions of α and γ particles, there could be a law like that governing the interactions of α and γ particles, even though those interactions never occur. All that it takes, on the universals account, is for there to be the right necessitation relations between universals.

7 The Problem of Induction, day 2

7.1 Review: Hume's First Problem

1. Last time we encountered David Hume, he was persuading us that the method by which we form beliefs about the future—enumerative induction—is not guaranteed to lead us to truth.

   (a) His reason, in a nutshell, was this: we can easily imagine all of the premises of an enumerative induction being true, but the conclusion being false. Thus, when we reason as follows:

   P1. The first fire caused smoke

   P2. The second fire caused smoke

   ⋮

   PN. The Nth fire caused smoke

   C. Fire will always cause smoke

   we can imagine the first N fires all causing smoke, but the $N+1$st fire not causing smoke. And, if we can imagine this, then it is possible. So it is possible that our reasoning leads us astray.

2. But perhaps Hume was wrong about the way we reason about these matters. Perhaps the inference we *really* draw is something like this:

   P1. The first fire caused smoke

   P2. The second fire caused smoke

   ⋮

   PN. The Nth fire caused smoke

   C1. It is a law that fire causes smoke

   C2. Fire will always cause smoke

(a) So: we began thinking about what a law of nature is, with an eye to thinking about how we might come to know what the laws of nature are.

(b) However, after considering those accounts of laws of nature, we can see that Hume's problem from part 1 of *Sceptical Doubts Concerning the Operations of the Understanding* remains. Just as we may conceive of P1, P2, . . . , and PN being true while C2 is false, so too may we conceive of P1, P2, . . . , and PN being true while C1 is false.

   i. On a 'Best Systems' understanding of laws, it could be that the regularities we have observed up to now will stop holding tomorrow—so that, from now on out, fires do not produce smoke. And, if that were so, then the best executive summary of facts from our world would not include or entail the statement that fire causes smoke. So it wouldn't be a law of nature that fire causes smoke. So our premises could be true while the conclusion is false.

   ii. On a Universals understanding of laws, it could be that fire *always* causes smoke, but accidentally so. That is, it could be that, even though fire always causes smoke, there is no necessitation relation between the universal of being a fire and the universal of producing smoke. So our premises could be true while the conclusion is false.

7.2 Hume's Second Problem

3. Okay, so perhaps induction isn't *guaranteed* to lead to truth. What of it? We may still have excellent *reason* to believe that induction will lead to truth. It's just that this reason isn't *conclusive* reason. Perhaps we will be led astray and fire will stop causing smoke, but this is very improbable, since we have very good reasons to think fire will continue to cause smoke.

   (a) Here's another example of enumerative induction:

   P1. The sun rose on the morning of the first day.

   P2. The sun rose on the morning of the second day.

   ⋮

   PN. The sun rose on the morning of the Nth day.

   C. The sun will rise tomorrow morning.

   (b) We may concede to Hume that it is *possible* that the sun won't rise tomorrow morning, but still recognize that we have plenty of good reasons to think that it will.

4. Hume's second problem—and, really, this is *the* problem of induction—is the denial of *this* claim. Hume doesn't think that you have *any* good reason to think that the sun will rise tomorrow.

(a) Why is that? Well, in order for you to reach the conclusion that the sun will rise tomorrow morning, Hume believes that you must assume that nature operates uniformly—what he calls the *principle of the uniformity of nature*. You have a good reason to think that the sun will rise tomorrow morning only if you have a good reason to think that nature operates uniformly.

(b) But, Hume contends, you *don't* have any good reason to think that nature operates uniformly. The only reason you could provide for thinking that this is true is that nature has always operated uniformly *in the past*:

P. Nature has always operated uniformly in the past

C. Nature will operate uniformly in the future

But *this is just another induction*. And, just as in our first induction, you have a good reason to accept the conclusion on the basis of the premises only if you assume that nature will always operate uniformly:

P. Nature has always operated uniformly in the past

UN. Nature operates uniformly

C. Nature will operate uniformly in the future

But our belief that nature operates uniformly is exactly what we were trying to justify. So we have just reasoned in a circle. And circular reasons are not good reasons. So we have no good reason to think that nature always operates uniformly.

(c) Let's put that all down in a nice premise-conclusion argument. Here is what Hume is saying:

P1. You have good reason to think that the sun will rise tomorrow only if you have good reason to think that nature operates uniformly.

P2. The only reason you have to think that nature operates uniformly is a circular reason.

P3. Circular reasons are not good reasons.

C1. You have no good reason to think that nature operates uniformly. (from P2 and P3)

C2 You have no good reason to think that the sun will rise tomorrow. (from P1 and C1)

5. Let's go through that again. Hume is interested in understanding the basis of your belief that the future will resemble the past. He thinks that, without some good reason to think that this is true, we have no reason to believe any of the predictions of our scientific theories. But, he thinks that the only reason we could possibly have for thinking that the future will resemble the past is that, *in the past*, the future has always resembled the past.

P1. The future has always resembled the past in the past

C. The future will resemble the past in the future

(a) But this argument is not deductively valid—the premise could be true while the conclusion is false. If we are going to have any good reason to think that the conclusion is true, on the basis of the premise, it must be because we are assuming that the past provides some guide to the future. That is, it must be because we are assuming that the future will resemble the past.

P1. The future has always resembled the past in the past

P2. The future will resemble the past

C. The future will resemble the past in the future

(b) But this argument is *blatantly* circular. We have just assumed the thing we are attempting to justify. Circular reasons aren't good reasons, so we don't have any good reason to think that the future will resemble the past.

6. To see Hume's problem more clearly, let's think about a competitor to induction, which we may call *counterinduction*.

(a) When the counterinductivist sees a regularity in nature, they predict that that regularity will *not* continue into the future.

(b) Thus, the counterinductivist reasons as follows:

P1. The sun rose on the morning of the first day.

P2. The sun rose on the morning of the second day.

⋮

PN. The sun rose on the morning of the *N*th day.

C. The sun will *not* rise tomorrow morning.

(c) When pressed on this inference, the counterinductivist patiently explains that they are relying upon a principle of *non-uniformity of nature*, that the future will not resemble the past:

P1. The sun rose on the morning of the first day.

P2. The sun rose on the morning of the second day.

⋮

PN. The sun rose on the morning of the *N*th day.

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

C. The sun will *not* rise tomorrow morning.

(d) 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, the principle *will* hold true in the future.

P1. The future has always resembled the past in the past

P2. The future will *not* resemble the past

C. The future will *not* resemble the past in the future

(e) It looks as though the counterinductivist's reasons for thinking that the sun won't rise tomorrow aren't very good. And it looks as though the *reason* that they are not very good is that those reasons are circular. But it also looks as though our own reasons for thinking that the sun *will* rise tomorrow are circular in the same way. So it looks as though we don't have any good reason to think that the sun will rise tomorrow.

8 | The Problem of Induction, day 3

1. When we're thinking about inductive inferences, here are two questions we might ask:

   Q1: Which inductive inferences are good, and which are bad?

   Q2: Why do we think that good inductive inference will lead us to truth?

   (a) Hume assumes that our answer to Q1 will be that the good inductions are those which assume that the future will resemble the past, and that nature will operate uniformly.

   (b) From this assumption, he argues that we will never be able to give a satisfactory answer to Q2. Our answers to Q2 will either be unsatisfactory, or they will be *circular*.

2. But it doesn't really matter what answer we give to Q1. Whatever answer we give, Hume is still going to be able to argue that we cannot give a satisfactory answer to Q2.

   (a) Here's why: if we're going to give an answer to Q2, then we're going to have to provide some *reasons*—let's just call those reasons, $p_1, p_2, p_3, \ldots$ (it doesn't really matter what the reasons are)—for thinking that induction will lead us to the truth.

$$\begin{array}{c} p_1 \\ p_2 \\ p_3 \\ \vdots \\ \hline \textbf{I}\text{: Induction will lead to truth} \end{array}$$

   (b) Then, Hume is going to want to know: do these reasons, $p_1, p_2, p_3, \ldots,$ *deductively* entail that induction will lead us to the truth? Or do they merely lend it *inductive* support?

What kind of support do p_1, p_2, and p_3 lend to **I**?

Deductive | Inductive

(c) Hume's dilemma is that, either way we go, there will be problems.

   i. We will never be able to support the conclusion **I** with premises which *deductively entail* that induction will lead us to the truth. For, if there were a deductively valid argument from known premises to the conclusion that induction will lead us to truth, then it would be *impossible* for those premises to be true while the conclusion **I** is false. But we can conceive of possibilities in which everything we know about the world is true, yet induction leads us to falsehood, and therefore, **I** is false.
   ii. On the other hand, if we attempt to justify the conclusion **I**, then our justification will be *circular*.
   iii. We should here distinguish two kinds of circular arguments: *premise* circular arguments and *rule* circular arguments.
   A *premise* circular argument contains its conclusion as one of its premises. For instance, the following is a premise circular argument:

      P1. We should trust Billy.

      P2. Billy says that we should trust him.

      C. We should trust Billy.

   On the other hand, a *rule* circular argument attempts to justify the use of a rule by *using* that very rule. For instance, the following arguments are *rule* circular:

      P1. Billy says that we should trust him.

      C. We should trust Billy.

      P1. If my logic professor isn't lying, then *modus ponens* is a valid rule of inference.[1]

      P2. My logic professor isn't lying.

      C. *Modus Ponens* is a valid rule of inference.

   Then, the second horn of Hume's dilemma is this: if we attempt to justify our assumption that induction will lead to truth *inductively*, then we will have provided a *rule circular* reason to think that induction will lead to truth.

3. (a) Thus, if our reason for thinking that induction will lead to truth is *deductive*, then it will fail to be deductively valid.
   (b) If, on the other hand, our reason for thinking that induction will lead to truth is *inductive*, then it will be rule circular.
   (c) Rule circular reasons are not good reasons, so we have no good reason to think that induction will lead to the truth.

[1] *Modus Ponens* is the rule of inference which says that, from a claim of the form 'If *P*, then *Q*', and a claim of the form '*P*', you may infer '*Q*'.

Part II

Falsification

9 Popper's Falsificationism, day 1

9.1 Popper's Solution to the Problem of Demarcation

1. Popper begins his essay by discussing the *problem of demarcation*.

   (a) To refresh your memory: many people think that there is an important difference between *sciences* like the fields on the left, and so-called *pseudo-sciences*, like those on the right.

Science	Pseudo-Science
Physics	Astrology
Medicine	Homeopathy
Chemistry	Phrenology
Biology	Freudian Psychology
Cosmology	Marxism

   The problem of demarcation is the problem of saying what it is that distinguishes the intellectual activities from the latter. In virtue of what are the intellectual activities of Physicists *scientific*, while the intellectual activities of Astrologers is not?

2. Here is a popular answer to the problem of demarcation (it is the answer given by Bacon and Hempel):

   (a) The fields on the left seek out *verification* of of their theories—they seek out evidence which could *confirm* their theories, in accordance with the inductive method. The theories they construct then *explain* a wide variety of data.

   (b) On the other hand, the fields on the right do not attempt to verify their hypotheses by seeking evidence which could confirm them. The theories they construct are not explanatory.

Science	Pseudo-Science
Seeks to verify its theories with evidence	Does not seek to verify its theories with evidence
Uses the inductive method	Does not use the inductive method
Has theories which are *explanatory*	Does not have theories which are *explanatory*

   (c) Call this the *verificationist* answer to the problem of demarcation.

3. Popper thinks that the verificationist's answer is wrong.

   (a) He illustrates why he thinks this answer is incorrect by considering three theories popular in the Vienna of his youth: Einstein's theory of gravitation, Marx's theory of history, and Freud's theory of the unconscious.

   (b) Popper notes that both Marx's and Freud's theories had *tons* of evidence verifying them. As soon as you adopt Marx and Freud's perspectives, you begin seeing evidence for them *everywhere*.

   (c) Similarly, both Marx's theory and Freud's theory could explain lots of phenomena. Marx's theory explained every major historical development—and Freud's theory explained tons of otherwise puzzling human behavior.

   (d) However, Popper began to think that this was actually a bad thing—their verification by evidence and their great explanatory power was a *weakness* of those theories, and not a strength.

   (e) Popper 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.

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

      i. While Marx and Freud's theories could easily accommodate and explain any evidence whatsoever, Einstein's theory would not have been able to accommodate or explain it if the apparent position of the stars were not changed during the solar eclipse.

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

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

   (a) On Popper's view, the goal of science is not to *verify* or *confirm* theories at all.

   (b) Rather, the goal of science is just to *falsify* theories.

Science	Pseudo-Science
Seeks to falsify its theories	Seeks to verify, and not to falsify, its theories

9.2 Popper's 'Solution' to the Problem of Induction

> *According to Popper, what makes a theory* corroborated*? Is this the same thing as a theory being* confirmed *or* verified*? Why or why not? Does Popper think that well-corroborated theories are more likely to be true than uncorroborated theories?*

5. Popper believes that his solution to the problem of demarcation provides us with a response to Hume's problem of induction, as well.

   (a) The response it affords us is this: while Hume is correct that induction is never justified, and never provides us with a good reason to think that a scientific hypothesis is true, this isn't a problem, because *science doesn't use induction*.

6. Recall: the verificationist uses the predictions of a theory to verify it, according to the following logic:

$$\frac{\text{If } H,\ \text{then } E \qquad E}{H}$$

   (a) *This* is an inductive inference. The truth of its conclusion is not guaranteed by the truth of its premises. So we face Hume's problem.

   However, a Popperian *falsificationist* never concludes that their theories are *true*. Rather, they only ever seek to *falsify* their theories. The only inference they ever perform is *this* one:

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

   (b) And *this* inference is deductively *valid*.

7. Since science, according to Popper, only ever seeks to *falsify* theories, and never seeks to *verify* or *confirm* them, science need only utilize deduction. It need not utilize induction at all.

   (a) So, the problem of induction does not arise.

10 | Popper's Falsificationism, day 2

1. Recall the distinction between *corroboration* and *confirmation*.
   (a) To say that a theory is *confirmed* by a test is to say that the outcome of the test makes the theory *more likely to be true*—that the outcome of the test has given us some *reason to believe* that the theory is true.
   (b) To say that a theory is *corroborated* by a test is only to say that the theory has not been falsified by the test. It is emphatically *not* to say that we have any reason to think that the theory is true. It is only to say that the theory survived the test.
2. Recall the distinction between two different attitudes you could take towards a scientific theory: *acceptance*, on the one hand, and *belief*, on the other.
   (a) To *accept* a theory is to use it for the purposes of prediction, to continue to subject it to test, and to otherwise organize your inquiry around the theory. In order to accept a theory, you need not believe that the theory is true (in fact, you could be firmly convinced that the theory is false).
   (b) To *believe* a theory is to think that what the theory says about reality is correct—it is to think that the theory is true.
      i. Note that you can *accept* a theory without *believing* it, and also that you can *believe* a theory without *accepting* it (by just thinking that it's true but still not using the theory for prediction, or to ever subject it to test, or otherwise organizing your inquiry around the theory).
      ii. For instance, an evolutionary biologist could *believe* in intelligent design without *accepting* it (because, for whatever reason, they don't think their religious beliefs should guide scientific inquiry). And a physicist could *accept* Quantum Field Theory without *believing* it (since it cannot explain gravity).
3. Popper thinks that tests of theories can *corroborate* hypotheses but never *confirm* them. He therefore thinks that scientists should *accept* their theories, but not *believe* them.
4. Let us begin by distinguishing two (separable) components of Popper's philosophy of science:

(a) *Rejectability*: in order for your inquiry to count as scientific, there must be some evidence you could receive which would lead you to *reject* the theory you currently accept. If you are not prepared to reject your theory in the light of any evidence, then your inquiry is not scientific.

(b) *Anti-Inductivism*: in order for your inquiry to count as scientific, you must only reason *deductively*. If you reason inductively, then your inquiry does not count as scientific.

5. Notice that we may accept *Rejectability* without accepting *Anti-Inductivism*.

   (a) For instance, you might think that you should believe the theory of general relativity because of the outcome of Eddington's observations. (That is, you might think that you should reason *inductively*.) However, you could still think that you should stop believing the theory of general relativity if its predictions are not borne out.

   (b) This is the attitude of Verificationists like Bacon and Hempel.

6. In the following, we will be raising some objections, not for Popper's thesis of *Rejectability*, but, instead, only for his *Anti-Inductivism*.

10.1 Problems with Anti-Inductivism

7. The first problem with *Anti-Inductivism* is that it appears to conflict with *Rejectability* when we are considering probabilistic hypotheses.

   (a) For instance, consider the hypothesis, H, that the chance that the coin lands heads, on any given flip, is 1/2.

   H: the chance that the coin lands heads is 1/2.

   (b) Suppose that we flip the coin 1,000,000 times and get the evidence E, that, every single time the coin is flipped, it lands heads.

   E: The coin lands heads 1,000,000 times in a row.

   (c) E is precisely the kind of evidence that would ordinarily lead us to *reject* the hypothesis H. However, if we were to reason *deductively*, then we would not be able to reject the hypothesis H on the basis of this evidence.

   (d) We might try to reject it by reasoning as follows:

   P1. If H, then it is not the case that E.

   P2. E.

   C. It is not the case that H.

   This argument is deductively valid; however, its first premise is false. The hypothesis H could be true, and even so, the coin could land heads every single time it is flipped. The hypothesis H assigns this outcome a positive (if incredibly tiny) probability.

(e) Recognizing this, we might try to reject the hypothesis H by reasoning as follows:

P1. If H, then probably it is not the case that E.

P2. E.

C. It is not the case that H.

While P1 is true, *this* argument is deductively invalid. Its premises could be true, all while its conclusion is false. For the hypothesis H could be true—the coin could be fair—even while the coin lands heads every time it's flipped. That's possible, which means that it's possible for the premises to be true while the conclusion is false. Which means that this argument is deductively invalid.

8. Probabilistic hypotheses like these appear throughout science.

   (a) The theory of Statistical Mechanics says only that entropy will increase *with very high probability*.

   (b) The theory of Quantum Mechanics doesn't predict any one experimental outcome—instead, it provides us with a probability distribution over possible outcomes.

   (c) In Evolutionary biology, claims are made about the *likelihood* of animals with certain characteristics reproducing.

9. These kinds of hypotheses appear to be paradigm instances of scientific theories. Popper's *anti-inductivism* says that they are not scientific. So it appears that Popper's *anti-inductivism* is incorrect.

P1 Anti-inductivism says that merely probabilistic hypotheses are not scientific.

P2 Merely probabilistic hypotheses are scientific.

C Anti-inductivism is false.

   (a) Note that the case of probabilistic hypotheses gives another example of a way in which we could accept the *rejectability* component of Popper's view without accepting the *anti-inductivism*.

      i. We could say that we would reject the hypothesis that the coin is fair, if it landed heads more than 95% of the time.

      ii. In so doing, we would be reasoning inductively.

10. Popper has a response to worries about probabilistic hypotheses. He says that

> *Probability statements, in so far as they are not falsifiable, are metaphysical and without empirical significance; and in so far as they are used as empirical statements they are used as falsifiable statements.*

   (a) The thought is this: even though the probabilistic hypothesis H isn't refutable with deductive reasoning alone, the hypothesis H^* *is* refutable by deductive reasoning alone:

H^*: the chance that the coin lands heads is 1/2 *and* in large trials, it will not land heads more than 95% of the time.

(b) This hypothesis, unlike H, may be deductively refuted by the evidence E, for the following argument is deductively valid, and has true premises:

P1 If H^*, then it is not the case that E

P2 E

C It is not the case that H^*

(c) So Popper contends that H^* is (because falsifiable) scientific, but H is (because not falsifiable) unscientific.

11. Probabilistic hypotheses like H (or even H^*) can also help us to illustrate another objection to Popper's anti-inductivism. The objection is that anti-inductivism requires us to foreswear beliefs which appear to be eminently rational.

(a) Suppose that we are testing the hypothesis H^*,

H^*: the chance that the coin lands heads is 1/2 *and* in large trials, it will not land heads more than 95% of the time.

and we receive the evidence that, in 1,000,000 flips, the coin landed heads exactly 500,000 times.

E^*: In 1,000,000 flips, the coin landed heads half of the time

(b) Popper's anti-inductivism means that we cannot conclude, on the basis of this evidence, that H^* is likely to be true.

P1 If H^*, then probably E^*

P2 E^*

C H^*

For this inference would be *inductive*.

12. If we think that we *should* believe that the coin is close to fair after receiving this evidence, then we may mount the following argument against Popper's anti-inductivism.

P1 Anti-inductivism says that we ought not believe that the coin is close to fair after seeing half of 1,000,000 flips land heads and half land tails.

P2 We should believe that the coin is close to fair after receiving this evidence.

C Anti-inductivism is false.

11 Popper's Falsificationism, day 3

1. Popper's Falsificationism is the conjunction of the following two claims about what makes an inquiry *scientific* (as opposed to *pseudo-scientific*):

   (a) *Rejectability*: in order for your inquiry to count as scientific, there must be some evidence you could receive which would lead you to *reject* the theory you currently accept. If you are not prepared to reject your theory in the light of any evidence, then your inquiry is not scientific.

   (b) *Anti-Inductivism*: in order for your inquiry to count as scientific, you must only reason *deductively*. If you reason inductively, then your inquiry does not count as scientific.

2. The *Duhem-Quine Thesis* says that scientific theories may not be tested on their own. Rather, in order for a scientific theory to be subjected to test, we must make a large number of additional *auxiliary assumptions*. It is only with these additional auxiliary assumptions that any test implications may be derived from the theory.

   (a) For instance, consider Newton's theory of Universal Gravitation. This theory tells us that all massive bodies exert a gravitational force upon each other, that force is proportional to the mass of the two bodies, and inversely proportional to the square of the distance between them. And it tells us that bodies will accelerate in the direction of the forces acting upon them, with a magnitude equal to the magnitude of the force acting upon them divided by the body's mass. That is, the theory provides us with the equations,[1]

$$F = ma$$

$$F^G_{1,2} = G \cdot \frac{m_1 \cdot m_2}{d^2_{1,2}}$$

   Now, suppose that we wish to use this theory to derive a test implication about the orbits of the planets. In order to do so, we must assume that the sun and the planets in our solar system are the only relevant masses—that all other masses are small enough or far enough away that their gravitational masses are negligible. We must also assume that the planets are *only* subject to the gravitational

[1] Here, G is some constant, $F^G_{1,2}$ is the force, due to gravity, which body 1 exerts upon body 2, m_1 is the mass of body 1, m_2 the mass of body 2, and $d_{1,2}$ is the distance from body 1 to body 2.

forces they exert upon each other. That is, we must assume that there are no electromagnetic forces (*e.g.*) which make a substantial difference with respect to the orbits of the planets.

More generally, *any* motion of the planets could be squared with Newton's theory by simply postulating additional forces.

(b) For another example, suppose that we have a weight loss drug and we wish to test the hypothesis that it speeds up the body's metabolism (or whatever). We could test this hypothesis by randomly splitting a selected group of people in two, giving half of them the drug and giving the other half a placebo. Then, we could measure the participant's weight after 3 months, 6 months, and 9 months to see what changes take place.

In order for our theory to yield the prediction that those who are assigned the drug will lose weight, we must make the following auxiliary assumptions: the people who are assigned the drug will actually remember to take it, and will actually take it with food, at the appropriate time. We need to assume that those on the placebo won't, because they are being weighed in this experiment, begin going to the gym more often. More generally, we need to assume that lifestyles will remain more or less constant throughout the study. We need to assume that the drug doesn't *also* cause an increase in hunger, or stress, or what-have-you, which could offset the increase in metabolic rate. We need to assume that our initial randomization didn't just by accident group all of the people whose weights were already on a downward trajectory into one group and all those whose weights were already on an upward trajectory into another group.

3. If the Duhem-Quine thesis is correct, then the simple schema for the logic of hypothesis testing we have been pre-supposing up to this point is incorrect. Recall, we have been supposing that, in testing a hypothesis H, we first derive a test implication from the hypothesis, E, and then check to see whether this test implication is true or false. If true, then the hypothesis is confirmed (or so say Bacon and Hempel, though Popper disagrees). If false, then the hypothesis is falsified.

Confirmation	Falsification
If H, then E	If H, then E
E	Not E
H	Not H

However, if the Duhem-Quine thesis is correct, then no hypothesis *on its own* generates any test implication whatsoever. It is only with the addition of *auxiliary assumption* that a hypothesis will make predictions. Then, the logic of hypothesis testing is more faithfully represented like so:

Confirmation*	Falsification*
If H and A, then E	If H and A, then E
E	Not E
H	Not H

4. This appears to pose a problem for Popper's *Anti-Inductivism*. The problem is that the argument schema Falsification* is an *inductive* argument (it is *not* deductively valid).

   (a) Popper's Falsificationism was supposed to solve Hume's problem of induction by demonstrating that science never used induction. He was able to do this because he supposed that science only ever *falsified* hypotheses, and he supposed that the logic of falsification was given by the schema Falsification, which is deductively valid. If, however, the logic of falsification is actually given by Falsification*, then, in falsifying hypotheses, we would no longer we reasoning *deductively*.

   (b) For instance, if we get the evidence that those in the treatment group didn't lose any weight at the end of the 9 months, we could always claim that the people in the treatment group didn't take the drug regularly, or with food. We could always claim that the people in the placebo group started eating less because of the study. There are any number of things would could say to square our hypothesis with the evidence we had received.

   Note: the Duhem-Quine Thesis *doesn't* tell us that doing this would be *rational* or *scientific*. However, it does tell us that, if we said these things, we wouldn't be describing an impossible scenario. So, it is possible for the premises of Falsification* to be true while its conclusion is false. So that argument schema is not deductively valid.

   (c) So, here is the problem for Popper's *Falsificationism*:

   P1 If the *Duhem-Quine Thesis* is correct, then scientific hypotheses may never be deductively refuted by any evidence.

   P2 If *Falsificationism* is correct, then scientific hypotheses must sometimes be deductively refuted by evidence.

   P3 The *Duhem-Quine Thesis* is correct.

   C If there are any scientific hypotheses, then *Falsificationism* is not correct.

5. Popper has a response to this worry. He writes:

> *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) It is unclear how this answers the argument above. In particular, it does not appear to rescue the Anti-Inductivist component of Popper's theory from the Duhem-Quine Thesis. If 'such a procedure' is always possible, then it appears as though we can never reason *deductively* from the evidence to the rejection of the hypothesis.

(b) But we could at least understand Popper here as defending the *Rejectability* component of his view. He says that, though you *could* claim that the weight loss drug works even after seeing nobody in the treatment group lose weight, doing so would be unscientific. In order for your inquiry to be scientific, you must be willing to reject your theory when its predictions are not borne out. And you must *not* attempt to hang onto your theory by replacing some of your auxiliary assumptions.

Scientific	Unscientific
If H and A, then E	If H and A, then E
Not E	Not E
Not H	Not A

6. Putnam has a response to this. He thinks that there are cases in which scientists responded to disconfirmatory evidence by rejecting their auxiliary assumptions, and not rejecting their theory. And he thinks, moreover, that these cases are paradigm examples of *good* scientific reasoning, and not unscientific pseudo-science, as Popper would have it.

   (a) In the early 19th century, Leverrier used Newton's theory of universal gravitation (**TUG**) to make predictions about the orbits of the planets in our solar system. In so doing, he made the following auxiliary assumption:

   A All objects other than the sun and the 7 planets are either small enough or distant enough that their gravitational influences are negligible.

   (b) However, the assumption **A**, together with the theory **TUG**, made false predictions (**U**) about the orbit of Uranus.

   (c) In response to this, Popper says that the scientific way to reason is as follows:

   P1 If **TUG** and **A**, then **U**

   P2 It is not the case that **U**

   C It is not the case that **TUG**

   (d) In fact, this is not what Leverrier did. He did not take Newton's theory to be rejected by the irregularities in Uranus's orbit. Rather, he decided to reject the auxiliary assumption **A**.

   P1 If **TUG** and **A**, then **U**

   P2 It is not the case that **U**

   C It is not the case that **A**

In its stead, he proposed a new auxiliary assumption: he proposed that there is an 8th planet, which he gave the name 'Neptune'. He then replaced the assumption **A** with the following assumption.

A′ All objects other than the sun and the 8 planets are either small enough or distant enough that their gravitational influences are negligible.

with this new assumption, **TUG** makes the correct prediction about the orbit of Uranus.

7. Putnam's point isn't just that Leverrier was *correct* about the existence of the planet Neptune. His point is that Leverrier's reasoning here was *good reasoning*, and, moreover, *scientific*.

   (a) Putnam therefore concludes that Popper's Falsificationism 1) does not match scientific practice, and 2) moreover, does not give good advice about how scientists *ought* to proceed.

Part III

Confirmation Theory

12 Confirmation Theory, day 1

12.1 Review

1. Bacon suggested that science operate through the application of an *algorithmic procedure* for generating and justifying scientific hypotheses.

   (a) Hempel argued that no such algorithmic procedure is possible in the *context of discovery*—though perhaps it is possible in the *context of justification*. Such an algorithmic procedure would provide a *logic of induction*.

2. Recall, we distinguished two different questions about inductive inferences:

   Q1: Which inductive inferences are good, and which are bad?

   Q2: Why should we think that good inductive inferences will lead to truth?

   Hume's *problem of induction* was that we have no good answer to Q2. The reasons we give will either deductively entail that good inductive inferences will lead to truth—and Hume takes himself to have shown that any such reasons will not be known—or else, the reasons we give will only *inductively support* that good inductive inferences will lead to truth—and such reasons will be *rule circular*.

3. We briefly considered whether rejecting a Humean conception of laws (like, *e.g.*, the Best Systems account) would allow us to escape Hume's problem. But it appears that the problem remains no matter how we think about laws of nature.

4. Popper offered us a way out— we needn't worry about the problem of induction, since science does not make use of induction at all. Science only seeks to *falsify* hypotheses, so it can make do with deduction alone.

5. Unfortunately, we saw powerful reasons to think that, on this score, Popper was wrong. Science does need induction, after all. For,

   (a) Deduction alone will not allow us to falsify merely probabilistic hypotheses.

   (b) The Duhem-Quine thesis tells us that no hypothesis may be deductively falsified by any test.

   (c) Putnam argued that some hypotheses shouldn't be rejected when their predictions are not borne out.

12.2 Back to the Problem of Induction

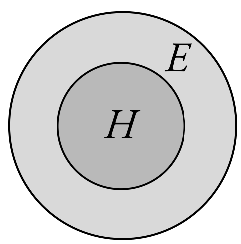
6. So: we've seen reasons to be skeptical of Popper's way out of the problem of induction. But showing that Popper's way out of the problem is no good is not yet to solve the problem. What should we say about the problem of induction?

   (a) Susan Haack makes the following observation: the problem of induction isn't only a problem for induction. It is a problem for deduction as well.

   (b) Suppose that we wish to justify the use of the deductively valid inference rule *Modus Ponens*,

   $$\frac{\text{If } H\text{, then } E \qquad H}{E}$$

   We could try to do so by noting that, if the first premise is true, then all the relevant ways H could be true are ways for E to be true, too. That is: the truth of H is *sufficient* for E—if we're in the inner circle, then we must be in the outer circle as well:

   And the second premise tells us that we are in the inner circle. So, we must be in the outer circle.

   (c) But wait—there's something suspicious about the reasoning we just engaged in. It appears to have the following form:

   If we're in the inner circle, then we're in the outer circle

   We're in the inner circle

   We're in the outer circle.

   But this argument utilizes the very rule we were attempting to justify! It is rule-circular in precisely the way that our justification of *induction* was rule circular.

   (d) Just as a counter-inductive rule of inference could justify itself, a counter-*deductive* rule could be used to justify itself, too. Suppose that we wish to justify the rule of inference that Haack (somewhat meanly) calls *Modus Morons*:

   $$\frac{\text{If } H\text{, then } E \qquad E}{H}$$

We could do so by making the following argument: from the first premise, we know that the truth of H is sufficient for the truth of E—that is, if we're in the inner circle above, then we must be in the outer circle as well. And the second premise tells us that we are in the outer circle. So, we must be in the inner circle, too.

(e) That is, somebody who defends *Modus Morons* could justify it with the following rule-circular argument:

If we're in the inner circle, then we're in the outer circle

We're in the outer circle

We're in the inner circle

7. The big picture point of Haack's example:

   (a) When we justify our *deductive* rules of inference, we must *use* those very rules of inference. And this is rule circular in precisely the way that an attempted justification of induction is.

   (b) Moreover, just as with counterinduction, there are other, *counterdeductive* rules of inference which may also be used to justify themselves.

   (c) So, Hume's problem of induction is not unique to induction. It arises for *de*duction as well. The probem of rule circularity is pervasive.

   (d) None of this, of course, helps us to *solve* Hume's problem of induction. It is, however, to point out that it is an instance of a far more general problem.

8. Two further points on Haack:

   (a) If Popper thinks that Hume's problem of induction gives us a reason to forego inductive inferences, then he should presumably think that Haack's problem of deduction also gives us a reason to forego *deductive* inferences. But then, his falsificationism cannot get off the ground.

   (b) Earlier in the semester, Hempel raised the possibility that we could give an *algorithmic* procedure for *justifying* hypotheses by the acquired evidence. That is, he raised the possibility of a logic of induction. We did not get excited about this possibility because of Hume's problem of induction—if we couldn't show that induction will lead to truth, then why care about how it works? However, now that we've seen that the very same problems plague *deduction*, which has a very well worked out formal logic, perhaps we should turn our attention back to the first question about induction that we put aside earlier in the semester, that is, the question Q1:

   Q1: Which inductive inferences are good, and which are bad?

12.3 Confirmation Theory

9. *Deductive* logic is the study of which deductive inferences are good deductive inferences (which are deductively valid), and which are bad deductive inferences (which

are deductively invalid). For instance, deductive logic will tell us that *Modus Ponens* (on the left) is deductively valid, whereas *Modus Morons* (on the right) is deductively invalid.

Modus Ponens	Modus Morons
If H, then E	If H, then E
H	E
———	———
E	H

10. An *in*ductive logic (or, as we'll call it—a *theory of confirmation*) would tell us which inductive inferences are good inductive inferences, and which are bad ones. That is, it would tell us that the inference

    All observed swans are white.
    ———
    All swans are white.

    is a good inductive inference, whereas the counter-inductive inference

    All observed swans are white.
    ———
    All unobserved swans are black.

    is a bad one.

11. (a) Sometimes a piece of evidence, E, gives us a reason to believe a hypothesis, H. When this is so, say that the evidence E *confirms* the hypothesis H.
    (b) Other times, a piece of evidence, E, gives us a reason to *dis*believe a hypothesis, H. When this is so, say that the evidence E *disconfirms* the hypothesis H.

12. Note that confirmation is not the same as *belief*. Just because we have some evidence that confirms H, that doesn't mean that we should think that H is true. We can see this more clearly by considering two points.
    (a) In the first place, confirmation comes in degrees. E could give a *very strong* reason to believe that H, or it could give just a rather *weak* reason to believe that H. If E confirms H, but only very weakly, then we shouldn't believe H on the basis of E alone.
    (b) In the second place, even if we have evidence that strongly confirms a hypothesis, that confirmation can be *defeated* by other evidence we have.
        i. For instance, let H be the hypothesis that John robbed the bank, and let E be the evidence that 100 eyewitnesses who know John personally identified John as the bank robber.
        ii. E strongly confirms H. But this doesn't necessarily mean that you should believe that H is true. Perhaps E isn't your *total* evidence.
        iii. Perhaps you additionally have the evidence that John has an identical twin brother, and John has a rock-solid alibi.
        iv. In that case, you shouldn't believe that H is true.

13. (a) In general, *deductive* inference is *indefeasible*—if P deductively entails C, then $P\&Q$ will *also* deductively entail C.
    (b) However, inductive inference is *defeasible*. It could be that E confirms H, even though $E\&F$ does not confirm H.
    (c) So: a theory of confirmation should only be expected to tell us whether, if E were our *total* evidence, we would have reason to believe H.

14. Ideally, what we would like from a theory of confirmation is this:
    (a) We would like that theory to tell us, for any given piece of evidence E, and hypothesis H, whether E confirms H, disconfirms H, or is evidentially neutral with respect to H.
    (b) It would additionally be nice if the theory could tell us, for any given piece of evidence E and any hypothesis H, to what *degree* E confirms or disconfirms H.
    (c) Finally, it would be nice if our theory of confirmation was—like our theory of deductive validity—both *formal* and *intersubjective*.
        i. A theory is *formal* if it looks at the syntax of E and H alone, and otherwise ignores the meaning of the statements. Deductive logic is formal because, once we know the *logical form* of a premise, we don't need to know anything else about what it means—*e.g.*, in "If H, then E", we don't need to know what "H" and "E" mean.
        ii. A theory is *intersubjective* when everybody can agree about whether E confirms H. It doesn't matter what else we believe about the way the world is, we'll agree about whether the total evidence E confirms the hypothesis H.

15. These are features we want a theory of confirmation to have. Unfortunately, when we meet again after the midterm, we'll discover that you can't always get what you want.

13 Confirmation Theory, day 2

13.1 Review

1. What we want from a theory of confirmation:

    (a) A *qualitative* account of confirmation.

    i. For any H, E: does E confirm H?

    (b) A *quantitative* measure of confirmation.

    i. For any H, E: to what *degree* does E confirm H?

    (c) We'd like our theory of confirmation to be both *formal* and *intersubjective*.

    i. *Formal*: we can say whether E confirms H by looking only at the syntax, or logical form, of E and H.

    ii. *Intersubjective*: we can all agree about whether E confirms H.

13.2 You can't always get what you want

2. Hempel: any theory of confirmation which satisfies these two plausible principles will say that every proposition confirms every other proposition

    Entailments Confirm (EC)
    If H entails E, then E confirms H.

    Consequence Condition (CC)
    If E confirms H, then E confirms anything which H entails.

    (a) Take any two propositions A and B.

    (b) $A \,\&\, B$ entails A. So, by **EC**, A confirms $A\&B$.

    (c) A confirms $A\&B$ (above) and $A\&B$ entails B. So, by **CC**, A confirms B

3. Hempel: even if we weaken these principles like this,

    Laws are Confirmed by Their Instances
    A law statement of the form "All Fs are Gs" is confirmed by an F G.

Equivalence Condition
If E confirms H, then E confirms anything which is logically equivalent to H.

nearly everything will end up confirming any universal law statement. Take, for a toy example, the law statement "All ravens are black".

(a) "All ravens are black" is logically equivalent to "All non-black things are non-ravens"

(b) By Laws are Confirmed by Their Instances, a green leaf confirms the hypothesis that all non-black things are non-ravens

(c) By (a), (b), and Equivalence Condition, a green leaf confirms the hypothesis that all ravens are black.

4. Goodman: in order to say whether "All Fs are Gs" is confirmed by an F G, we must know something about what 'F' and 'G' *mean*.

   (a) Say that a thing is grue iff it has been observed before 2020 and is green or has not been observed before 2020 and is blue.

   Grue
   Something is grue iff it is first observed prior to 2020 and is green, or it is not first observed prior to 2020 and is blue.

   (b) Then, there is no *syntactic, formal* difference between *this* inductive inference (which is a strong inductive inference):

   The first observed emerald is green
   The second observed emerald is green
   ⋮
   The *n*th observed emerald is green
   ———
   All unobserved emeralds are green

   and *this* inductive inference (which is a counter-inductive inference):

   The first observed emerald is grue
   The second observed emerald is grue
   ⋮
   The *n*th observed emerald is grue
   ———
   All unobserved emeralds are grue

5. A purely formal theory of confirmation cannot distinguish induction from counterinduction. So a theory of confirmation must go beyond logical form.

13.3 The New Riddle of Induction

6. A lingering question: what *is* the difference between the hypothesis that all emeralds are green, and the hypothesis that all emeralds are grue?

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

(b) Response: Well, yes, if you start off talking in terms of 'green' and 'blue', then 'grue' will be defined relative to a time. But, suppose that you start off talking in terms of 'grue' and 'bleen', where

    Bleen
    Something is bleen iff it is first observed before 2020 and is blue or it is not first observed before 2020 and is green.

    Then, if I were to try to explain to you the meaning of 'green' and 'blue' in terms you would understand, I would have to make reference to specific times. In particular, I would have to give you the following definitions:

    Green
    Something is green iff it is first observed before 2020 and is grue, or it is not first observed before 2020 and is bleen.

    Blue
    Something is blue iff it is first observed before 2020 and is bleen, or it is not first observed before 2020 and is grue.

(c) So, for those who begin speaking in terms of 'green' and 'blue', the properties of grueness and bleenness will appear to be complicated gerrymandered properties defined relative to a time. But, similarly, for those who begin speaking in terms of 'grue' and 'bleen', the properties of greenness and blueness will appear to be complicated gerrymandered properties defined relative to a time.

(d) Is there anything we could say to the grue-speakers to make them recognize that something has gone wrong with their (counter?)inductive inferences?

14 Confirmation Theory, day 3

14.1 Review

1. Last time, we saw two reasons to think that we can't get everything we might want out of a theory of confirmation—in particular, we saw two reasons to think that our theory cannot be purely formal.

   (a) Two seemingly reasonable formal conditions we laid down—that Laws are Confirmed by their Instances and the Equivalence condition—entail that *almost everything* confirms a hypothesis. In particular, a white piece of chalk confirms the hypothesis that all ravens are black. (Hempel's *Ravens Paradox*)

   (b) If our theory is purely formal, then it will not be able to distinguish the properties of *greenness* and *grueness*. But a theory of confirmation should distinguish these two, since a collection of green/grue emeralds confirms the hypothesis that all emeralds are green, but not the hypothesis that all emeralds are grue.

14.2 A Probabilistic Theory of Confirmation

2. Here's a promising thought: what a purely formal theory of confirmation has left out is the *probability* of a hypothesis like H, given the evidence E.

3. Given a suitable probability function, we may say:

   (a) If the probability of H, *given that*, or *conditional on*, E, is greater than the unconditional probability of H,

   $$\Pr(H \mid E) > \Pr(H)$$

   then E confirms H

   (b) If the probability of H, *given that*, or *conditional on*, E, is less than the unconditional probability of H,

   $$\Pr(H \mid E) < \Pr(H)$$

   then E disconfirms H.

(c) And, if the probability of H, *given that*, or *conditional on*, E, is equal to the unconditional probability of H,

$$\Pr(H \mid E) = \Pr(H)$$

then E neither confirms nor disconfirms H.

4. In order to better understand this theory, let us walk through a (brief) introduction to the theory of probability.

14.3 The Theory of Probability

5. A *probability function* Pr is a function from *propositions* to real numbers between 0 and 1 inclusive.

(a) If Pr maps a proposition, A, to the number x, then we write

$$\Pr(A) = x$$

6. We will characterize probability functions by the rules that they obey. Any probability function will satisfy the following rules:

The 'Not' Rule For any proposition A,

$$\Pr(\textit{not } A) = 1 - \Pr(A)$$

(a) For instance, to find the probability that a fair die *doesn't* land on 1 (not 1), we calculate

$$\begin{aligned} \Pr(\textit{not } 1) &= 1 - \Pr(1) \\ &= 1 - \frac{1}{6} \\ &= \frac{5}{6} \end{aligned}$$

The 'Or' Rule For any propositions A and B,

$$\Pr(A \textit{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \textit{ and } B)$$

(a) For instance, suppose that we flip a fair coin twice. The probability that it lands heads *both* times is 1/4. To find the probability that it lands heads *either* on the first flip *or* on the second flip (that is, the probability that it lands heads at least once), we calculate

$$\begin{aligned}
\Pr(H_1 \textit{ or } H_2) &= \Pr(H_1) + \Pr(H_2) - \Pr(H_1 \textit{ and } H_2) \\
&= \frac{1}{2} + \frac{1}{2} - \frac{1}{4} \\
&= \frac{3}{4}
\end{aligned}$$

The Equivalence Rule For any logically equivalent propositions A and B,

$$\Pr(A) = \Pr(B)$$

(a) For instance, since $A\&B$ is logically equivalent to $B\&A$,

$$\Pr(A\&B) = \Pr(B\&A)$$

(b) And, since 'All ravens are black' is logically equivalent to 'All nonblack things are nonravens',

$$\Pr(\textit{ All ravens are black }) = \Pr(\textit{ All nonblack things are nonravens })$$

7. Any probability function can be thought of as a *muddy Venn diagram*.

   (a) You take a Venn diagram and slap some mud on it. The proportion of the total mud sitting on top of any proposition is the probability that that proposition is true.

8. Let's introduce the following stipulative definitions,[1]

Conditional Probability For any propositions A and B, the *conditional* probability of A, given B, $\Pr(A \mid B)$, is

$$\Pr(A \mid B) \stackrel{\text{def}}{=} \frac{\Pr(A\&B)}{\Pr(B)}$$

Probabilistic Independence The propositions A and B are *independent* iff

$$\Pr(A\&B) = \Pr(A) \cdot \Pr(B)$$

[1] If $\Pr(B) = 0$, then $\Pr(A \mid B)$ will be undefined.

(a) It follows from these two definitions that A and B are probabilistically independent iff

$$\begin{aligned} & \Pr(A \mid B) = \Pr(A) \\ \text{and} \qquad & \Pr(B \mid A) = \Pr(B) \end{aligned}$$

9. It is often easier to think in terms of conditional probabilities than unconditional probabilities. With this definition in hand, we may formulate the following additional probability rules.

> **The 'And' Rule** For any statements A and B,
>
> $$\Pr(A \& B) = \Pr(A \mid B) \cdot \Pr(B)$$

(a) For instance, suppose that I have a an urn with 8 purple marbles and 2 red marbles. And suppose that I randomly draw two marbles from the urn, one after the other, without replacing the first before selecting the second. We may calculate the probability that I draw two red marbles as follows.

$$\begin{aligned} \Pr(\text{Red}_1 \& \text{Red}_2) &= \Pr(\text{Red}_2 \mid \text{Red}_1) \cdot \Pr(\text{Red}_1) \\ &= \frac{1}{9} \cdot \frac{2}{10} \\ &= \frac{2}{90} \\ &= \frac{1}{45} \end{aligned}$$

> **The Rule of Total Probability** For any statements A and B,
>
> $$\Pr(A) = \Pr(A \mid B) \cdot \Pr(B) \; + \; \Pr(A \mid \textit{not } B) \cdot \Pr(\textit{not } B)$$

(a) For instance, suppose that there is a disease which 10% of people have. There's a test for the disease which is 90% reliable—that is, given that you have the disease, D, the probability that the test says you have the disease, T, is 90%, and, given that you don't have the disease, *notD*, the probability that the test says you have the disease is 10%. What is the probability that the test says that you have the disease?

$$\begin{aligned} \Pr(T) &= \Pr(T \mid D) \cdot \Pr(D) \; + \; \Pr(T \mid \text{not } D) \cdot \Pr(\text{not } D) \\ &= \frac{9}{10} \cdot \frac{1}{10} \; + \; \frac{1}{10} \cdot \frac{9}{10} \\ &= \frac{9}{100} \; + \; \frac{9}{100} \\ &= \frac{18}{100} \end{aligned}$$

Bayes' Rule (v1) For any statements A and B,

$$\Pr(A \mid B) = \frac{\Pr(B \mid A) \cdot \Pr(A)}{\Pr(B)}$$

(a) For instance, given that the test above *says* that you have the disease, the probability that you actually do is 50%, or 1/2, since

$$\begin{aligned} \Pr(D \mid T) &= \frac{\Pr(T \mid D) \cdot \Pr(D)}{\Pr(T)} \\ &= \frac{9/10 \ \cdot \ 1/10}{18/100} \\ &= \frac{9/100}{18/100} \\ &= \frac{9}{18} \\ &= \frac{1}{2} \end{aligned}$$

10. Combining **Bayes' Rule (v1)** with the **Rule of Total Probability**, we get

Bayes' Rule (v2) For any statements A and B,

$$\Pr(A \mid B) = \frac{\Pr(B \mid A) \cdot \Pr(A)}{\Pr(B \mid A) \cdot \Pr(A) \ + \ \Pr(B \mid \text{not } A) \cdot \Pr(\text{not } A)}$$

15 Confirmation Theory, day 4

15.1 Bayesian Confirmation Theory

1. The Bayesian interprets probability claims as being about the *degrees of belief*, or the *credences*, C, of some rational person (the scientist).
   (a) If $C(A) = 1$, then the scientist thinks that A is certainly true.
   (b) If $C(A) = 0$, then the scientist thinks that A is certainly false.
   (c) If $C(A) = 1/2$, then the scientist is as confident that A is true as they are that A is false.
2. The Bayesian then endorses the following norms of rationality:

   Probabilism
   The scientist's credence function C is (or should be, or is representable as being) a *probability function*.
   Conditionalization
   Upon acquiring the total evidence E, the scientist should be disposed to adopt a new credence function C_E which is their old credence function *conditionalized on* E. That is, for all H,

$$C_E(H) = C(H \mid E) = \frac{C(E \mid H)}{C(E)} \cdot C(H)$$

3. The Bayesian theory of confirmation says that E confirms H iff

$$C_E(H) > C(H)$$

   And E disconfirms H iff

$$C_E(H) < C(H)$$

   (a) Terminology:
      i. Pr is the agent's *prior* credence function.
      ii. $\Pr_E$ is the agent's *posterior* credence function.
   (b) Notice the following consequence of the Bayesian theory of confirmation: E confirms H iff H makes E more likely than not-H does. That is,

Good Predictions Confirm E confirms H,

$$C_E(H) > C(H)$$

if and only if H makes E more likely than not-H does,

$$C(E \mid H) > C(E \mid \textit{not-}H)$$

(c) *Cf.* Popper—a hypothesis gets more confirmation the bolder its predictions. If a hypothesis makes *everything* likely, then it does not receive any confirmation.

15.2 Why the Bayesian Thinks You Can't Always Get What You Want

4. Recall: these two principles jointly entail that "All ravens are black" is entailed by a non-black non-raven.

   Laws are Confirmed by Their Instances
   A law statement of the form "All Fs are Gs" is confirmed by an F G.

   Equivalence Condition
   If E confirms H, then E confirms anything which is equivalent to H.

5. Bayesian confirmation theory rejects that Laws are Confirmed by their Instances. (Question: do they accept Equivalence Condition?)

   (a) A toy model: suppose that you are certain that there are 8 things in existence, and you split your credence equally between these two hypotheses about their properties:

   <u>All</u>

	Black	Non-Black
Raven	4	0
Non-Raven	2	2

   <u>Some</u>

	Black	Non-Black
Raven	2	2
Non-Raven	2	2

   i. You get the evidence E = a randomly selected thing is a non-black non-raven.
   ii. According to the Bayesian theory of confirmation, E will confirm *All* iff *All* makes E more likely than *Some* does. But

   $$\Pr(E \mid \textit{All}) = 1/4 \qquad \text{and} \qquad \Pr(E \mid \textit{Some}) = 1/4$$

   iii. So the hypothesis *All* is not confirmed by a non-black non-raven.
   iv. So Laws are Confirmed by Their Instances is false.

   (b) Contrast this with a case where you get the evidence E^* = a randomly selected thing is a black raven.

i. According to the Bayesian theory of confirmation, E^* will confirm *All* iff *All* makes E more likely than *Some* does. And

$$\Pr(E^* \mid All) = 1/2 \qquad \text{and} \qquad \Pr(E^* \mid Some) = 1/4$$

ii. So a black raven confirms *All*, even though a non-black non-raven does not.

15.3 Objections to Bayesian Confirmation Theory

6. Recall: a non-formal theory of confirmation cannot distinguish the Green hypothesis from the Grue hypothesis.

   (a) Notation:

      i. $Green$ = All emeralds are green
      ii. $Grue$ = All emeralds are grue
      iii. E = All observed emeralds are green/grue

   (b) It turns out that the Bayesian can only say that *Green* is more likely than *Grue*, given the evidence E, if *Green started out* more likely than *Grue* in the prior. For

$$\frac{\Pr(Green \mid E)}{\Pr(Grue \mid E)} = \frac{\frac{\Pr(E|Green)}{\Pr(E)} \cdot \Pr(Green)}{\frac{\Pr(E|Grue)}{\Pr(E)} \cdot \Pr(Grue)} = \frac{\Pr(E \mid Green) \cdot \Pr(Green)}{\Pr(E \mid Grue) \cdot \Pr(Grue)} = \frac{\Pr(Green)}{\Pr(Grue)}$$

   (c) Therefore, it will be the case that *Green* is more likely than *Grue after* receiving the evidence, $C_E(Green) > C_E(Grue)$, iff it *was* the case that *Green* was more likely than *Grue before* receiving the evidence.

   (d) It seems like the prior which gives higher credence to the hypothesis that all emeralds are green is better than the one which gives higher credence to the hypothesis that all emeralds are grue

   (e) But Bayesian Confirmation Theory doesn't tell us which priors to begin with

   (f) This is the "problem of the priors"—it is the problem of specifying which prior credences are reasonable.

7. Here is another objection to Bayesian Confirmation Theory:

   (a) The "perihelion" of Mercury is the point in its orbit when it is closest to the Sun.

(b) It was well-known that the perihelion of Mercury precesses; but this couldn't be explained by Newtonian celestial mechanics.

(c) It *could*, however, be explained by Einstein's Theory of General Relativity (TGR). This was widely regarded as *confirmation* of the TGR.

(d) But, if the precession of the perihelion of Mercury (E) was already well known, then scientists should have *already* conditionalized on it. And

$$C_E(E) = C(E \mid E) = 1$$

(e) But then, the precession of the perihelion of Mercury could not *confirm* TGR, according to Bayesian confirmation theory, since

$$C_E(TGR \mid E) = C_E(TGR)$$

(f) This is the problem of *old evidence*. The Bayesian cannot make sense of old evidence confirming a theory. But there seem to be many such examples from the history of science.

16 | Objective Chance

16.1 Objective Chance

1. Scientific theories are filled with probabilistic claims which are not understood as in any way encapsulating our *ignorance*, or anything about our *epistemic state*.

   (a) These are not *subjective* probabilities; rather, they are *objective* probabilities, or *chances*.

2. For example:

   (a) In Casinos, they want dice which are *fair*—dice which have an equal probability of any side landing up. They don't care about whether anyone has any good *reason to think* that a die will land 1-up. Even if nobody knows that the die is biased, the Casino will still lose money if the die *is* biased.

   (b) The half life of carbon-14 is 5.730 years. That is to say: the chance of an atom of carbon-14 decaying in 5,730 years is 1/2. This doesn't appear to be a claim about anybody's subjective degrees of belief. It appears to be a claim about the carbon-14 atom itself.

   (c) Quantum mechanics tells us that, if we observe the z-spin of an electron in an eigenstate of x-spin up, there is a 1/2 chance that it will have z-spin up and a 1/2 chance that it will have z-spin down. Again, it is difficult to understand this as making a claim about our subjective degrees of belief. How could our subjective degrees of belief have anything to do with the z-spin of an electron out in the world?

   (d) Benford's law tells us that you are more likely to find numbers beginning with the digit 1 than with the digit 9. The law holds up with data like GDP growth, fundamental constants of nature, length of rivers, and so on. Again, difficult to understand this as having anything to do with anyone's subjective degrees of belief.

3. Our question for this part of the course is this: if objective chances aren't subjective degrees of belief, then what *are* they?

4. One good general rule for investigating questions like this is to figure out what properties objective chances have, and then look for something which has these general

properties—that is, go looking for something which is well-suited to play the objective chance role. So, here are some general properties of objective chances:

(a) Consider a certain kind of chance trial known as a *Bernoulli* trial, where there are two possible outcomes with probabilities p and $1 - p$. A coin flip is a nice example—a flip has two possible outcomes, heads and tails, each with a probability of approximately 1/2. A the number of trials goes to infinity, the *relative frequency* of heads landings,

$$\frac{\#\text{ heads}}{\#\text{ flips}}$$

will *very likely* approach 1/2.

   i. Of course, in any finite number of flips, the relative frequency of heads landings may fall short of the chance; but you should expect the relative frequency to get closer and closer to the chance, as the number of flips gets larger and larger.

(b) We generally take objective chances to *explain why* the relative frequency converges to a certain value.

(c) Knowledge of the objective chances can rationally constrain our subjective probabilities, or credences.

   i. If we are certain that the chance of heads in 1/2, then our degree of belief that the coin lands heads on any given flip should be 1/2.

16.2 The Classical Account of Objective Chance

5. The *classical* account of objective chances says that facts about objective chance are *a priori*, and have to do with certain sorts of *symmetries*, and facts about how we should proportion our degrees of belief in states of *ignorance*.

6. On this view, when we say that the chance that the coin lands heads is 1/2, we are saying something about the symmetry of the coin and our lack of evidence about how it will land.

7. The classical account assigns probabilities in accordance with the **Principle of Indifference**.

   Principle of Indifference Outcomes which are evidentially symmetric have equal probabilities. If we have no more reason to believe that p than we have reason to believe that q, then the probability of p should equal the probability of q.

8. For instance, the chance that a die lands 1-up is 1/6, since we have no more reason to believe that it will land 1-up than we have reason to believe that it will land 2-up, 3-up, 4-up, 5-up, or 6-up. Applying the **Principle of Indifferences**, each of these outcomes should have an equal probability.

9. A problem for the Classical Account:

Cross-Country Drive (v1)
I drove 2100 miles from Pittsburgh to L.A. The trip took somewhere between 30 and 42 hours. What is the chance that it took between 30 and 35 hours?

(a) Applying the principle of indifference, we get that $\Pr(30 \leqslant t \leqslant 35) = 5/12$.

Cross-Country Drive (v2)
I drove 2100 miles from Pittsburgh to L.A. My average velocity was somewhere between 50 and 70 mph. What is the chance that the average velocity was between 60 and 70 mph?

(a) Applying the principle of indifference, we get that $\Pr(60 \leqslant v \leqslant 70) = 1/2$.

But these aren't two different cases. They are the very same case, described in two different ways. A 2100 mile trip will take somewhere between 30 and 42 hours if and only if my average velocity is somewhere between 50 and 70 miles per hour. And it will take between 30 and 35 hours if and only if my average velocity is between 60 and 70 miles per hour. So the **Equivalence Rule** from probability theory tells us that we must assign these two propositions the same probability.

10. So it looks as though the principle of indifference leads to contradictions.

16.3 Actual Frequentism

11. According to the *actual frequentist*—sometimes called the *finite frequentist*—the objective chances just are the actual relative frequencies.

    (a) Unlike the classical account, this account says that the objective chances are *a posteriori*.

    (b) *Cf.* regularity theories of laws of nature.

12. More carefully, the actual frequentist says that:

    Actual Frequentism the objective chance of an outcome O (in a reference class K) is

$$Ch_K(O) = \frac{\text{\# of } K\text{s which are } O}{\text{\# of } K\text{s}}$$

    (a) For instance, to see the probability that a coin lands heads (within the reference class of flipped coins), you count up how many times a flipped coin has landed heads and divide that by the total number of times a coin has been flipped. That ratio is the chance that a flipped coin lands heads.

13. Note that, according to actual frequentism, claims about chances only make sense given some reference class.

(a) My chance of contracting cancer could therefore be different depending upon whether we're talking about my chance of getting cancer as a 33-year old man, or as a man, or as former smoker living in Pittsburgh, or,...

14. This gives rise to the following objection to actual frequentism: suppose that there are only two coins—a double-headed coin and a double-tailed coin. Each day we flip them both. Then, what is the objective chance that the double-headed coin lands heads on a particular flip?

    (a) The actual frequentist says: that depends upon the reference class. If we group it in the reference class of *coin flips*, then the chance that it lands heads will be 1/2. If we group it in a reference class containing one of the double-headed flips and 3 of the double-tailed flips, then the chance that it lands heads will be 1/4. With other reference classes, we can get the chance that it lands heads to be whatever we wish.

    (b) But it doesn't appear that the chance that this coin lands heads should depend upon how we group it together in reference classes in this way.

    (c) Since this coin is double-headed, it appears that the chance that this coin lands heads should be 1.

    (d) Actual frequentism says otherwise, so actual frequentism is wrong.

15. Here's another objection to actual frequentism (the problem of the single case):

    (a) Consider a possible world in which only a single coin (with a heads face and a tails face) is ever minted. That coin is flipped once and immediately destroyed.

    (b) Actual frequentism says that the chance that this coin lands heads is 1, if it actually does land heads, and 0, if it doesn't land heads.

    (c) But it seems as though this coin could very well be a fair coin—with equal chance of landing heads and landing tails.

    (d) Actual frequentism says that it could not be a fair coin.

    (e) So, actual frequentism is wrong.

16. Another objection to actual frequentism.

    (a) A number is *rational* if it can be written as a fraction, a/b. Not all numbers are rational (in fact, in a good sense *most* numbers are *ir*rational).

    (b) According to actual frequentism, an objective chance *must* be rational. This looks like a problem for actual frequentism, since, in many scientific theories, there are irrational chances—that is, there are objective probabilities $\Pr(A) = x$, where x cannot be expressed as a ratio a/b.

17. A related worry:

    (a) According to actual frequentism, it is impossible for a fair coin to be flipped an odd number of times.

(b) For, if the coin is only ever flipped an odd number of times, then the number of times it lands heads cannot be 1/2 of the number of times it is flipped.

18. One final objection to actual frequentism, mirroring an objection we raised to regularity theories of laws of nature earlier in the semester:

    (a) Objective chances should *explain* why the actual frequencies are what they are.

    (b) But, if objective chances *just are* the actual frequencies, then they cannot *explain* those frequencies.

Part IV

Scientific Realism

17 Scientific Realism, day 1

1. Recall the distinction between *metaphysics* and *epistemology*.
   (a) Metaphysics is the branch of philosophy which asks questions about which things exist, and how they exist.
      i. Are there laws of nature? If so, what kind of things are they? Are they just certain kinds of regularities, or patterns? Or are they instead relations between universals which *govern* and *explain* the regularities and patterns? Or are they something else altogether?
      ii. Are there objective chances? If so, what kind of things are they? Are they certain *a priori* symmetries? Or are they *a posteriori* frequencies? Or are they something else altogether?
   (b) Epistemology is the branch of philosophy which asks questions about whether/how we may come to know things about the world, and whether our beliefs are justified or not.
      i. Do we know (are we justified in believing) that the future will resemble the past?
      ii. Do we know (are we justified in believing) that all emeralds are green, as opposed to grue?
      iii. Are we justified in being more confident that all ravens are black after seeing a non-black, non-raven?
2. Let us add an additional kind of philosophical question to this list: questions about *semantics*, or the *meaning* of our language.
   (a) A *semantic* question is a question about the *meaning* of our language. For instance, when we say "The coin will likely not land heads on every flip", what is it that we *mean*?
   (b) Or, when I say "I know that the bank is open", what do I *mean*?
   (c) Or, when I say "there are electrons being shot out of a cathode ray inside of the television set", does the word "electron" successfully refer to something out in the world?
3. There will be important connections between these three kinds of questions.

(a) For instance, if the world is a certain way (metaphysics)—for instance, if it is filled with misleading evidence, as in the Matrix—then it becomes difficult to see how we could know that the world is this way (epistemology).

(b) If the things we claim to know are about our own sensory experience (semantics), then it is easy to see how we could come to know these things (epistemology).

(c) If we have a theory of meaning according to which we can only every talk *about* our own sense experience (metaphysics), then many metaphysical questions will begin to seem *meaningless*.

4. But the questions are still importantly different, and it's important that we keep them clearly separated when thinking about philosophical questions.

17.1 Eddington's Two Tables

5. You look at a table. While looking at the table, you have a certain visual experience of the table. Pointing, you say "there's a table". Here's a natural, common-sensical understanding of what's going on:

   (a) Your experience of the table is distinct from the table itself. (After all, you and I have different experiences of the table, but our different experiences are of the very same table.)

   (b) When you say "table", your word is talking *about* the table, and not the experience—that is, the word "table" successfully *refers* to the table. (After all, when you point and say "there's a table", you and I are talking about the very same thing; but we have two different experiences, so if we were talking about our experiences, we would be talking about two different things.)

6. A brief terminological aside: let's say that a term *refers* to some thing out in the world if it attempts to talk about that thing. And let's say that a term *successfully refers* to some thing out in the world if it *succeeds* in talking about that thing.

   (a) For instance, in the sentence "Santa Claus is probably exhausted after delivering presents all around the world",

      i. "probably" does not refer to anything in the world (probably)

      ii. "Santa Claus" refers, but not successfully

      iii. "the world" both refers and successfully refers.

7. Then, our natural, common-sensical understanding of what's going on consists of the following three claims (one metaphysical, one epistemic, and one semantic)

   (a) The table exists, and it exists independently of our experience of it

   (b) We know that the table exists by observing it

(c) When we say "table" we successfully refer to that mind-independent thing in the world—the table.

8. Eddington worried that the picture of the world presented to us by mature science seems to put pressure on this naive understanding of the world and our interaction with it.

> One of [my tables] has been familiar to me from earliest years. It is a commonplace object of that environment which I call the world...It has extension; it is comparatively permanent; it is coloured; above all it is substantial...Table No. 2 is my scientific table. It is a more recent acquaintance and I do not feel so familiar with it. It does not belong to the world previously mentioned—that world which spontaneously appears around me when I open my eyes...My scientific table is mostly emptiness. Sparsely scattered in that emptiness are numerous electric charges rushing about with great speed; but their combined bulk amounts to less than a billionth of the bulk of the table itself.

9. This begins to put pressure on the naive, common-sensical understanding of the world and our relation to it.

    (a) For instance, when we say that the table is brown, the naive understanding would be that we are talking about properties of the mind-independent thing—the table, the thing out there in the world.

    (b) But if what's really out there in reality is Eddington's *scientific* table, then it does not appear to be colored at all. So what are we talking about when we say that the table is brown? Are just mistaken?

    (c) There is some pressure to start thinking that what we are really talking *about* is just our own appearances—that when we say that the table is brown, we are really only talking about our *experience* of the table.

10. These considerations give rise to metaphysical, epistemic, and semantic questions:

    (a) Do colors exist? If so, what are they? Are they intersubjective?

    (b) Can we know which colors things have? If so, how?

    (c) What are our color terms—like "brown"—really talking about? Are they talking about the table, or are they talking about our experience of the table, or something else altogether? Do they succeed in talking about anything at all?

11. Tables and colors are at least *observable*. Matters seem to get more complicated when we turn from the odd scientific duplicates of our everyday experience to the scientific entities which appear nowhere in everyday experience—entities which are not observable at all.

    (a) Science seems to tell us that the images on our television screens (or at least, the images on the old television screens from my youth) are created by a beam consisting of tons of tiny particles called electrons being shot out of a cathode ray tube at the back of the television. The beam of electrons moves so quick that we perceive a static image.

(b) It seems to tell us that our visual impressions are impressions of those electrons.

(c) It seems to tell us that when we say the word "electron", we successfully refer to those tiny unobservable particles.

17.2 Scientific Realism

12. The view known as *scientific realism* accepts each of these claims. That is, the *scientific realist* accept the following three claims:

    (a) A semantic claim: the theoretical terms appearing in our best scientific theories refer to mind-independent entities.

    (b) A metaphysical claim: they *successfully* refer; those mind-independent entities really exist, and the things we believe about them are true.

    (c) An epistemological claim: we know that these mind-independent entities really exist, and we know that the things we believe about them are true.

18 Scientific Realism, day 2

1. Recall, the scientific realist accepts three claims:
   (a) A semantic claim: the theoretical terms appearing in our scientific theories refer to mind-independent entities.
   (b) A metaphysical claim: the *successfully* refer; that is, those mind-independent entities really exist.
   (c) An epistemological claim: we know the things our theories tell us about those mind-independent entities.
2. Why should we accept scientific realism? One argument contends that, roughly, if scientific realism were false, then the predictive success of science would be miraculous—since there are 'no miracles', scientific realism must be true. This argument is known as the 'no miracles' argument for scientific realism.

18.1 The "No Miracles" Argument for Scientific Realism

3. This argument makes an *inference* to the best explanation.
   (a) What is an inference to the best explanation? It considers several hypotheses, and *infers* to the truth of whichever hypothesis provides the *best explanation* of our total evidence.
      i. For instance, suppose that, whenever you see lightning, you hear thunder. Here are three hypotheses which purport to explain this evidence: H_1: the lighting is causing the thunder. H_2: the clouds are causing both the lighting and the thunder. H_3: the lightning and the thunder are totally unrelated; they just happen to occur simultaneously.
      ii. H_1 provides a *better* explanation of the evidence than do H_2 or H_3. So we infer to the truth of H_1.
4. The evidence to be explained in the "no miracles" argument is the predictive success of modern physics. The best explanation is purported to be scientific realism.

P1. Every time physics makes a new prediction about the observable effects of electrons, we observe precisely those predictions.

P2. The best explanation of this is that electrons exist and that our theories about them are true.

C. Electrons exist and our theories about them are true.

(a) More generally, we may present the "no miracles" argument as:

P1. Mature scientific theories are wildly successful in making predictions in novel applications.

P2. The best explanation of this success is that the theoretical terms of our mature scientific theories successfully refer, and that the theory's claims are at least approximately true.

C. Mature scientific theories are at least approximately true.

18.2 The Underdetermination Argument Against Scientific Realism

5. Our best scientific theories are *underdetermined* by the evidence we have. For instance:

   (a) Curve-fitting. Given any finite collection of observations of X and Y, there are an infinite number of curves passing through the values of X and Y you've observed. So the true curve is *underdetermined* by the observations.

6. The underdetermination in the case of curve-fitting is an example of *weak underdetermination*.

   The Weak Underdetermination Thesis
   For any given theory T, and any observable evidence E which T predicts, there is some other theory T^* which also predicts E.

   (a) Even though there will always been an infinite number of curves consistent with any finite data set, we could always in principle figure out whether one or both of two such curves is the true curve by just collecting more data.

7. According to the *strong* underdetermination thesis, there are other, more radical, kinds of underdetermination.

   The Strong Underdetermination Thesis
   For any given theory T, there is some other theory T^* which makes *all the same observable predictions* as T does.

   If T and T^* make all the same observable predictions, then we will say that they are *empircally equivalent*.

8. Some purported examples of empirically equivalent theories:

   (a) Newtonian mechanics with the assumption that the center of mass of the universe is at rest, $TN(0)$, and Newtonian mechanics with the assumption that the center of mass of the universe is moving at a constant speed v, $TN(v)$.
   (b) There are several different interpretations of (non-relativistic) Quantum Mechanics, each of which may make the same *observable* predictions, but disagree about what the unobservable aspects of the world are like.
   (c) According to Einstein's general theory of relativity (GTR), spacetime is curved, and not Euclidean. However, by suitably modifying the theory, we may get another theory which makes all the same predictions as GTR, but according to which spacetime is Euclidean but there are several, new, distorting forces which stretch out and contract our measuring apparatus.

9. With either the weak or the strong underdetermination thesis, we may construct the following argument against scientific realism:

   P1. The Weak (Strong) Underdetermination Thesis

   P2. If the actual (any possible) observable evidence does not distinguish between T and T^*, then you do not have any reason to believe T rather than T^*.

   P3. If you do not have any reason to believe T rather than T^*, then you cannot know that T is true and that T^* is false.

   C. We do not know that our best scientific theories are true.

10. This conclusion contradicts the epistemological claim of scientific realism.

    (a) It thereby gives us a reason to reject the metaphysical claim, too. For, if we don't know that our theories are true, & if we have no reason to believe them, then we should not believe them. So we shouldn't accept the metaphysical claim of scientific realism.

11. One objection to this argument: P2 is false, for *empirical* success is not the only reason to favor one theory over another.

    (a) There are additionally *theoretical virtues* like simplicity (Ockham's razor).
    (b) These theoretical virtues may give us a reason to favor one of a collection of empirically equivalent theories.

12. Another objection to the argument: the strong underdetermination thesis is false. If T and T^* are empirically equivalent, then they are one and the same theory.

    (a) Suppose that somebody tried to argue for the strong underdetermination thesis by bringing up a theory according to which electrons are *positively* charged, and protons are *negatively* charged, and pointing out that this theory makes precisely the same predictions as standard theories about electricity and magnetism.

(b) A quick reply would be to point out that which charge we call positive and which we call negative is just a matter of *convention*. So the proposed theory is not really a *competitor* to the standard theory. It is rather a notational variant. The two theories are *physically equivalent*.

(c) Similarly, a *conventionalist* may wish to respond to the putative examples of empirical equivalence by saying that:

   i. $TN(0)$ and $TN(v)$ are physically equivalent; they are the same theory, and the choice between them is a matter of convention.

   ii. Theories with curved spacetime and theories with Euclidean spacetime and distorting forces are physically equivalent; they are the same theory, and the choice between them is a matter of convention.

   iii. Different interpretations of Quantum Mechanics are physically equivalent; they are the same theory, and the choice between them is a matter of convention.

13. Question: does conventionalism help the scientific realist?

   (a) In what sense could we say that 'curved spacetime' successfully refers, if we accept this response?

   (b) In what sense could we say that curved spacetime exists, if we accept this response?

   (c) in what sense could we be said to know that spacetime is curved, if we accept this response?

19 Logical Positivism

1. Recall, the scientific realist accepts three claims:
   (a) A semantic claim: the theoretical terms appearing in our scientific theories refer to mind-independent entities.
   (b) A metaphysical claim: they *successfully* refer; that is, those mind-independent entities really exist.
   (c) An epistemological claim: we know the things our theories tell us about those mind-independent entities.
2. We've seen the so-called 'no miracles' argument *for* scientific realism. We've also seen the underdetermination argument *against* scientific realism. This argument targeted the *epistemological* component of scientific realism.
3. There is another kind of scientific anti-realism—one which targets the *semantic* component of scientific realism, and denies that the theoretical terms appearing in our scientific theories refer to mind-independent entities.

19.1 Logical Positivism

19.1.1 The Verificationist Criterion of Meaningfulness

4. Distinguish the following two kinds of empiricism:
   (a) Epistemological Empiricism, which says that our knowledge of the world is derived from sense experience alone (it does not additionally derive from a faculty of pure *reason*).
   (b) Semantic Empiricism, which says that our *talk* about the world is really only talk about sense experience (we are not additionally talking about some mind-independent entities).
5. The Logical positivists were *semantic empiricists*. They denied that we can meaningfully talk about mind-independent entities—we can only ever meaningfully talk about our own sensory experience.

6. To understand some of the motivation for this view: the positivists were increasingly disenfranchised with the kind of philosophy which was prevalent around Vienna (where most of the logical positivists were based)—in particular, they were upset with Hegel, whose writing they found obscure and meaningless. The logical positivist Hans Reichenbach quotes the following passage from Hegel to illustrate:

    > *Reason is substance, as well as infinite power, its own infinite material underlying all the natural and spiritual life; as also the infinite form, that which sets the material in motion.*

    (a) Reichenbach didn't think that passages like this were false—they weren't *even* false, for they don't succeed in saying anything at all. Passages like these are, according to the positivists, simply *meaningless*.

7. The logical positivists wanted to lay down a criterion for meaningfulness—a criterion which would allow them to dismiss Hegelian claims like this as meaningless. The criterion they settled upon was this: in order to be meaningful, a statement must be *verifiable*—that is, in order to be meaningful, there must be some sense experience you could have which would *confirm* or *disconfirm* the statement. Else, the statement is meaningless.

    > The Verificationist Criterion of Meaningfulness (v1)
    > A statement is meaningful iff there is some way of *empirically testing* the statement to determine whether or not it is true.

    That is: a statement says something iff it says something about *sense experience*. A few comments on this criterion of meaningfulness.

    (a) Here, the empirical test need not be a *conclusive* test. It could be a test which would merely *confirm* or *disconfirm* the statement to some small degree. But, if the statement is going to mean anything at all, then it had better mean something about the kinds of sense experiences you could expect to have in various circumstances, were the statement true.
    (b) A statement can be meaningful and false at the same time—for instance, 'Hillary won the election' is meaningful (since there are ways of empirically testing it, by, *e.g.*, turning on the news), but false.

8. Notice that the following claims are, according to the Verificationist Criterion of Meaningfulness, simply meaningless:

    (a) Murder is wrong.
    (b) God exists.
    (c) God doesn't exist.
    (d) Properties inhere in substances.
    (e) There are no substances, merely properties.

    So not just Hegel, but much of traditional ethics and metaphysics, was ruled meaningless by the logical positivists.

19.1.2 Analytic and Synthetic Statements

9. An objection to **Verificationist Criterion of Meaningfulness (v1)**: what about claims like '$2 + 2 = 4$', or 'Every even number greater than 4 is the sum of two primes?' There is no way of *empirically testing* these claims—they don't say anything about sense experience.

   (a) The logical positivists acknolwedged this point. In response, they wanted to draw a distinction between two kinds of statements: *analytic* statements and *synthetic* statements.

      i. An analytic statement is one which can be known to be true or false simply by knowing the meanings of the words appearing in it. For instance, the logical positivists thought that statements like the following were analytic (the first two are analytically *true*, the second two analytically *false*):

         A. All bachelors are unmarried.
         B. If something is red, then it is colored.
         C. There is a mountain not surrounded by lower ground.
         D. Some woman is taller than herself.

      ii. A synthetic statement, in contrast, is one which can only be known to be true or false by knowing *both* the meaning of the words appearing in it *and* the way the world actually is. For instance, the logical positivists thought that statements like the following were synthetic:

         A. Some bachelors are bald.
         B. All apples are red.
         C. There is a mountain made of gold.
         D. No building is taller than the Empire State.

   (b) With the distinction between analytic and synthetic statements in hand, the logical positivists reformulate their verificationist criterion of meaning as follows:

      The Verificationist Criterion of Meaningfulness (v2)
      A **synthetic** statement is meaningful iff there is some way of *empirically testing* the statement to determine whether or not it is true.

      That is: a *synthetic* statement says something iff it says something about experience.

19.1.3 Observational and Theoretical Terms

10. What about the following claims—do *they* say something about experience?

   (a) Electrons are negatively charged.
   (b) The wavefunction evolves according to Schrödinger's Equation.
   (c) Hydrogen atoms have a single proton and a single electron.

At first, it's not clear how they could, since electrons, protons, and wavefunctions do not appear in our sensory experience.

11. Distinguish *observational* terms from *theoretical* terms. An observational term is a term which *directly* describes some possible sense experience. For instance, the following are all observational terms:

    (a) Red
    (b) Hard
    (c) Smelly
    (d) Trump

    A theoretical term, in contrast, is a term from a scientific theory which does not directly describe some possible sense experience. For instance, the following are all theoretical terms:

    (a) Charge
    (b) Charm
    (c) Quark
    (d) Gene

12. Similarly, an observation *statement* is just a statement which contains only observational terms. And a theoretical *statement* is just a statement which contains some theoretical terms. (Thus, any theoretical terms is enough to make a statement theoretical—in order to be observational, the statement must lack *any* non-observational terms.)

13. Then, the logical positivists claimed that, while theoretical terms do not *directly* describe experience, they do *indirectly* describe it. That is: they claimed that theoretical statements could be reduced, in some way or other, to purely observational statements. Thus, talking about unobservable entities like quarks is really just an oblique way of talking about possible experience.

    (a) For instance, to say 'there are electrons in the cloud chamber' is just to say that you will see cloud trails in the chamber.

14. A. J. Ayer attempted to carry this project out. Here is one of his first attempts:

    > Ayer's Verificationist Criterion of Meaningfulness (v1)
    > A synthetic statement is meaningful iff it is verifiable. And a statement is verifiable iff it entails an observation statement.

    (a) But wait—recall the Duhem-Quine Thesis. No scientific theory on its own will entail an observation statement. Auxiliary assumptions will be required.

    Very well—let us emend the criterion, then.

Ayer's Verificationist Criterion of Meaningfulness (v2)
A synthetic statement is meaningful iff it is verifiable. And a statement is verifiable iff it (perhaps together with some auxiliary statements) entails an observation statement (and the observation statement does not follow from the auxiliary statements alone).

(a) But wait—consider a paradigm case of a meaningless statement, 'Reason is infinite power'. This statement, together with the auxiliary statement:

    If Reason is infinite power, then the sofa is red

will entail the observation statement 'the sofa is red'. And 'the sofa is red' is not entailed by 'If Reason is infinite power, then the sofa is red' alone. So Ayer's second version of the verificationist criterion of meaningfulness says that 'Reason is infinite power' is meaningful.

(b) More generally, *for any statement* S, S together with

    If S, then the sofa is red

will entail the observation statement 'the sofa is red'. So *every* statement is meaningful, according to this revised account.

15. Ayer's verificationist criterion of meaning underwent further refinements, which were subject to further objections, and so on. (I'll spare you the details). Ultimately, Ayer decided that verificationism would have to await the development of an inductive logic.

    (a) Consider: you now know more than Ayer did about the history of philosophical investigations into inductive logic. Can you see a way of formulating the verificationist criterion of meaningfulness that avoids the kinds of objections raised above?

19.1.4 One Final Objection

16. A final objection to the verificationist criterion of meaningfulness (in any version): what test of the verificationist criterion of meaningfulness could we perform?

    (a) If none, then does the verificationist criterion of meaningfulness say that it itself is meaningless?

20 Constructive Empiricism

1. Recall, the scientific realist accepts three claims:
   (a) A semantic claim: the theoretical terms appearing in our scientific theories refer to mind-independent entities.
   (b) A metaphysical claim: they *successfully* refer; that is, those mind-independent entities really exist.
   (c) An epistemological claim: we know the things our theories tell us about those mind-independent entities.
2. We've see an argument for scientific realism—the so-called 'no miracles' argument—as well as an argument against the epistemological claim of scientific realism—this was the underdetermination argument.
   (a) Recall, the underdetermination argument came in two flavors—the *weak* underdetermination argument and the *strong* underdetermination argument.
3. The logical positivists rejected the *semantic* claim of scientific realism.
   (a) Given their alternative semantic claim—the verificationist criterion of meaningfulness, they took the metaphysical and epistemological claims of scientific realism to be meaningless.
4. Today we will be learning about a different kind of scientific anti-realism—Bas van Fraassen's *Constructive Empiricism*.

20.1 Constructive Empiricism

5. Unlike the logical positivists, van Fraassen accepts the semantic claim of scientific realism. He believes that our best scientific theories are referring to mind-independent entities. However, he denies that we are in a position to know that our best scientific theories are *successfully* referring to those entities. Nor can we know that what our theories say about those entities is true or even approximately true.
6. Nevertheless, van Fraassen agrees with the scientific realist that we should *accept* our best scientific theories. However, the scientific realist and van Fraassen will disagree about what it means to *accept* a scientific theory.

(a) For the scientific realist, to *accept* a scientific theory is to believe that it is *true*.

(b) For van Fraassen, to *accept* a scientific theory is to believe that it is *empirically adequate*—that is, it is to believe everything the theory has to say about *observable reality*.

7. This disagreement between van Fraassen and the realist stems from a disagreement about the aim of science.

   (a) According to the realist, the aim of science is to give us *true* theories.

      i. Thus, to accept a theory is to believe that it succeeds at its goal, and that it is true.

   (b) According to van Fraassen, the aim of science is to give us *empirically adequate* theories.

      i. Thus, to accept a theory is to believe that it succeeds at its goal, and that it is empirically adequate.

20.1.1 van Fraassen's Reply to the 'No Miracles' Argument

8. Recall, the scientific realist argued for their position with the 'no miracles' argument, which was an inference to the best explanation. The argument went as follows:

   P1. Mature scientific theories are wildly successful in making predictions in novel applications.

   P2. The best explanation of this success is that the theoretical terms of our mature scientific theories successfully refer, and that the theory's claims are at least approximately true.

   C. Mature scientific theories are at least approximately true.

9. van Fraassen has two replies to this argument. In the first place, he notes that a constructive empiricist like himself will object to the use of *inference to the best explanation*. After all, inference to the best explanation would allow us to infer to the existence of unobservable entities of our theories, supposing that our theories offer the best explanation of the data.

   (a) What the constructive empiricist should accept is not the rule of inference *infer to the truth of the best explanation*, but rather *infer to the empirical adequacy of the best explanation*.

   (b) But note that, if we try to re-cast the "no miracles" argument by inferring to the empirical adequacy of the best explanation, we get the following argument:

P1. Mature scientific theories are wildly successful in making predictions in novel applications.

P2. The best explanation of this success is that the theoretical terms of our mature scientific theories successfully refer, and that the theory's claims are at least approximately true.

C. Mature scientific theories are at least approximately **empirically adequate**.

(c) But van Fraassen is happy to accept *this* conclusion, since he thinks that we should accept mature scientific theories—that is, we should believe that they are empirically adequate.

10. van Fraassen's second reply to the "no miracles" argument concedes, for the sake of argument, the use of inference to the best explanation. However, he objects that the *truth* of those theories is not the best explanation of their empirical success.

    (a) There is another, better, explanation of the empirical success of our best scientific theories—an explanation in terms of *natural selection*:

    > *Any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive—the ones which* in fact *latched on to actual regularities in nature.*

20.1.2 van Fraassen on Supra-empirical Theoretical Virtues

11. Recall, one objection to the underdetermination argument against scientific realism appealed to theoretical virtues like simplicity (Ockham's razor) to distinguish between empirically equivalent theories.

    (a) van Fraassen accepts that simplicity is a good reason to prefer a theory—however, he thinks that it is only a *pragmatic* reason to favor a theory. A simpler theory will be easier to use.

    (b) However, just because a theory is easier to use than a rival, this does not give us any reason to think that it is any more likely to be *true* than its rival.

20.1.3 On Observability

12. van Fraassen's philosophy of science requires a sharp distinction between the *observable* and the *unobservable*. For van Fraassen only thinks that we should accept the *observable* consequences of our best-confirmed theories.

13. However, there is no sharp boundary between the observable and the unobservable. Rather, there is just a continuum. For we can observe with the naked eye, through glasses, through microscopes, through electron microscopes, through scanning tunneling microscopes, and so on.

(a) In fact, with a scanning tunneling microscope, we can 'observe' (?) individual atoms.

14. Noting this, Grover Maxwell objects that, since there's no clear dividing line between the observable and the unobservable, we will have to make a choice about where to draw this line—but why should we think that where we draw this line makes any difference with respect to what exists and what does not?

    (a) In response, van Fraassen makes two points: firstly, the difference made in where we draw this line is not between what exists and what does not (a metaphysical distinction), but rather between what we have *reason to believe* exists and what we don't have reason to believe exists (an epistemological distinction).

    (b) Secondly, van Fraassen notes that, just because the dividing line between the observable and the unobservable is vague, that doesn't mean that we are not entitled to the distinction.

        i. Compare: the dividing line between being bald and being not bald is vague, but that doesn't mean that Dmitri isn't bald. Though the boundary is vague, we are still entitled to distinguish between the bald and the non-bald.

15. How should we understand the notion of 'observability' used by van Fraassen? He tells us that a thing is *observable* iff "there are circumstances such that, if the thing is present to us under those circumstances, then we observe *it*."

    (a) Notice that, according to this definition, we don't have to *actually observe X* in order for *X* to be observ*able*. So distant planets will count as observable, as will Dodo birds—for, if we were transported through space and time so that these objects were before us, we would observe them.

    (b) So 'observable' means 'observable *in principle*', even if not in practice.

16. What about a creature that has electron microscopes for eyes? For such a creature, wouldn't electrons count as observable?

    (a) van Fraassen: perhaps, but when I accept the observable consequences of a theory, I merely accept the consequences about what *we* could in principle observe, not what *other creatures* could in principle observe.

    (b) So 'observable' means 'observable in principle *for us*'.

        i. Objection: the colorblind cannot even in principle observe colors; should they not believe in colors?

        ii. van Fraassen: the colorblind are a part of our epistemic community. So colors count as observable for *us*, and the colorblind should believe in them.

17. In sum: for van Fraassen, accepting a scientific theory means accepting everything it says about what members of our epistemic community could observe, were they in the right circumstances (even if they definitely won't be in those circumstances).

    (a) Note that, for van Fraassen, it is important that this notion of observability plays an important *epistemic* role.

(b) That is, van Fraassen thinks that we can know that distant, unobserved planets are attracted to massive bodies; however, we *cannot* know that electrons attract one another.

(c) Similarly, he thinks that we can know that colors exist, even if we are color blind, since we can rely upon the testimony of the color-sighted; however, we *cannot* know that DNA exists. We cannot rely upon the 'testimony' of electron microscopes.

20.1.4 Is van Fraassen selectively skeptical in an unmotivated way?

18. An objection to constructive empiricism: why should the difference between entities which are observable, given impossible changes in our spatio-temporal location and entities which are observable, given impossible changes to our perceptual capacities make a difference with respect to what we are justified in believing, given our *actual* evidence?

19. Another way of making this objection: notice that, while the *strong* underdetermination argument could be used to argue for van Fraassen's position, the *weak* underdetermination argument could not.

    (a) Recall: the weak underdetermination argument relies upon the weak underdetermination thesis:

    Weak Underdetermination Thesis For any given theory T, and any observable evidence E which T predicts, there is some other theory T^* which also predicts E.

    But this thesis says that the *empirical adequacy* of our best scientific theories is underdetermined by our actual evidence. So it pushes against accepting that our theories are empirically adequate just as much as it pushes against accepting that our theories are *true*.

    (b) So van Fraassen needs a response to the weak underdetermination argument—for it could be used to reject his view.

    (c) But then, van Fraassen faces the challenge of explaining why we should deny the premise

    > If the actual observable evidence does not distinguish between T and T^*, then you do not have any reason to believe T over T^*.

    even though we *accept* the premise

    > If *all possible* observable evidence does not distinguish between T and T^*, then you do not have any reason to believe T over T^*.

Part V

Scientific Explanation

21 Scientific Explanation, day 1

21.1 Scientific Explanation

1. Science doesn't just make *predictions* about the world; it additionally *explains* various facts about the world. For instance:

   (a) Why does the figure skater spin faster when they draw in their arms? Because angular momentum, L, is conserved, and $L = I\omega$, where ω is their angular velocity, and I is their *moment of inertia*. I is proportional to mass and the radius of rotation. So, when the figure skater brings in their arms, the radius of rotation decreases, and I goes down. Since L is constant, this means that their angular velocity, ω, must increase. So they spin faster.

   (b) Why did the marble settle at the bottom of the bowl? Because that's the only place in the bowl where the potential energy is a local minimum, and points where potential energy achieve a local minimum are stable equilibria towards which the marble will move. So the marble will end up at this equillibrium at the bottom of the bowl.

   (c) Why can't a vacuum raise water higher than 34 ft? Because the vacuum raises water by allowing the atmospheric pressure to push the water *outside* the vacuum down, and the atmospheric pressure only exerts a force sufficient to raise the water 34 ft.

2. These explanations seem like good explanations—however, some explanations seem bad. For instance, Francesco Sizi explained why there are exactly seven planets by citing the fact that there are exactly seven "windows" in the human head (the two eyes, two ears, two nostrils, and the mouth). This explanation is not as good as the others.

3. The goal of a philosophical theory of explanation is to say something about why the first three explanations are good, but the final explanation is bad.

   (a) That is, our question is this: in virtue of what do some facts provide good explanations of other facts?

21.2 The Deductive Nomological Account of Scientific Explanation

4. Hempel and Oppenheim (in 1948), conjectured that what makes an explanation good is that the *explanans* show that the *explanandum* was *to be expected*, and that it was to be expected as a matter of law.
   (a) Terminology: the *explanandum* is the fact to be explained in an explanation, and the *explanans* are the facts which do the explaining.
5. Hempel and Oppenheim develop this basic thought along the following lines: a good scientific explanation should be understood as a *deductively valid argument*, where the premises are the explanans, the conclusion is the explanandum, and one of the explanans includes a statement of law.
   (a) This account is known as the *deductive nomological* account—"nomological" means 'having to do with laws of nature'. The term comes from the Greek *νομοσ* (*nomos*), which means 'law'.
6. Schematically, a deductive nomological explanation (or just a 'DN Explanation') will take the following form:

$$\left.\begin{array}{c} C_1, C_2, \ldots, C_N \\ L_1, L_2, \ldots, L_M \end{array}\right\} \text{Explanans}$$
$$\overline{\qquad E \qquad} \Big\} \text{Explanandum}$$

   (a) where $C_1, C_2, \ldots, C_N$ are particular matters of fact, and $L_1, L_2, \ldots, L_M$ are universal generalizations.
7. A DN explanation will be good only if:
   (a) The argument is deductively valid; and
   (b) The premises are true.
8. Hempel and Oppenheim place further restrictions on a DN explanation. To motivate those additional restrictions, consider the following putative explanations:

Sam's son has measles.

Jim is Sam's son.

Jim has measles.

   (a) This does not explain why Jim has measles. So, Hempel and Oppenheim conclude, a DN explanation *must* include a statement of law—the law statement is not optional.

(b) Moreover, the law statement must be *necessary* for the argument to be deductively valid. We cannot transform the bad explanation above into a good explanation of why Jim has measles by simply including the additional premise that $F = ma$.

$F = ma$

Sam's son has measles.

Jim is Sam's son.

Jim has measles.

(c) In contrast, Hempel & Oppenheim allow an explanation to lack any statements of particular matters of fact, $C_1, C_2, \ldots, C_N$. For instance, the following is good explanation of why your bicycle doesn't travel faster than the speed of light:

Nothing travels faster than the speed of light.

Your bicycle didn't travel faster than the speed of light.

9. Hempel and Oppenheim additionally worry about explanations like these:

Massive bodies love each other.

Bodies which love each other attract each other.

The moon and the earth have mass.

The moon and the earth attract each other.

(a) They rule such explanations out by saying that the explanans must have *empirical content*.

10. Thus, we have the following account of what makes an explanation of a particular matter of fact a good one:

The DN Account A good scientific explanation is one in which:

(a) The explanans deductively entail the explanandum
(b) The explanans are all true
(c) The explanans have empirical content
(d) The explanans contain at least one universal generalization which is a law
(e) Without any given law statement, the remaining explanans do not deductively entail the explanandum

11. Let's call a putative explanation which satisfies all of these criteria a 'DN Explanation'.

12. Then, Hempel and Oppenheim claim that being a DN Explanation is both necessary and sufficient for being a good explanation (of a particular fact by a universal law).

22 Scientific Explanation, day 2

22.1 The DN Account

1. Recall, Hempel and Oppenheim gave the following account of what makes an explanation of a particular matter of fact a good one:

   The DN Account A good scientific explanation is one in which:

   (a) The explanans deductively entail the explanandum
   (b) The explanans are all true
   (c) The explanans have empirical content
   (d) The explanans contain at least one universal generalization which is a law
   (e) Without any given law statement, the remaining explanans do not deductively entail the explanandum

2. Let's call a putative explanation which satisfies all of these criteria a 'DN Explanation'.

3. Then, Hempel and Oppenheim claim that being a DN Explanation is both necessary and sufficient for being a good explanation (of a particular fact by a universal law).

22.1.1 Probabilistic Explanations

4. Note: we can sometimes explain phenomena by appealing to the facts that make it highly likely that the explanandum event would occur.

   (a) E.g., why has about 1/2 of the carbon-14 in this fossil decayed? Because the half-life of carbon-14 is 5,730 years, and the fossil has been around for about 5,730 years.

5. This is not a problem for Hempel and Oppenheim, for this is an explanation of a particular fact by a *statistical law*. But Hempel and Oppenheim are only giving an account of scientific explanations of particular facts by *universal laws*. They have a separate account of when statistical laws can explain particular facts.

6. In general, Hempel and Oppenheim recognize 4 different kinds of scientific explanations: explanations of particular facts by appeal to universal laws, explanations of regularities by appeal to universal laws, explanations of particular facts by statistical laws, and explanation of regularities by statistical laws.

Laws	Explananda: Particular Facts	Explananda: Regularities
Universal Laws	Deductive-Nomological	Deductive-Nomological
Statistical Laws	Inductive-Statistical	Deductive-Statistical

The DN Account we have provided above applies only to the first of these categories.

22.2 Objections to the DN Account

7. Consider the following putative explanation:

   Either Jim has measles or $F \neq ma$.

   $F = ma$.

   Jim has measles.

   (a) This meets all of the criteria laid down by the DN Account, but appears to be a terrible explanation of why Jim has measles.

8. Alternatively, consider this putative explanation:

   No one who takes birth control pills gets pregnant.

   Mary takes birth control pills.

   Mary doesn't get pregnant.

   (a) This meets all of the criteria laid down by the DN account, but it's not clear whether this is the correct explanation of why Mary didn't get pregnant. After all, it could be that Mary is a virgin. To twist the knife, notice that the explanation continues to satisfy all of the DN Account's criteria even if we exchange Mary with a biological male, Jim:

   No one who takes birth control pills gets pregnant.

   Jim takes birth control pills.

   Jim doesn't get pregnant.

   (b) But this is a terrible explanation. Jim's taking birth control pills doesn't have anything to do with his failure to get pregnant.

9. Finally, consider the following explanation of why the flagpole casts a shadow l meters in length:

   The sun is at an elevation of $\theta°$ in the sky.

   Light propagates linearly.

   The flagpole is h meters high.

   The length of the shadow is $l = h/\tan\theta$ meters.

(a) This appears to be a good explanation, and it meets all of the criteria laid down by the DN Account, but notice that it will *continue* to meet those criteria if we swap the third premise and the conclusion:

The sun is at an elevation of θ° in the sky.

Light propagates linearly.

The length of the shadow is $l = h/\tan\theta$ meters.

The flagpole is h meters high.

(b) But, while the height of the flagpole explains why the shadow is as long as it is, the length of the shadow does *not* explain why the shadow is as long as it is.

22.3 The Missing Ingredient

10. In most of the problems for the DN account, there isn't the right kind of *causal* relationship between the explanans and the explanandum. For instance:

    (a) It follows from the law that people who take birth control will not get pregnant, and that Jim took birth control, that Jim will not get pregnant. However, even so, Jim's taking birth control didn't *cause* him to not get pregnant.

    (b) The length of the shadow does not *cause* the height of the flagpole. It is the other way around—the height of the flagpole causes the length of the shadow.

11. This suggests the following hypothesis: what's missing from the DN account is a requirement that there be the right kind of *causal connection* between the explanans and the explanandum.

    (a) That is: it suggests that to explain a particular event is, in some way or other, to provide information about the *causes* of that event.

12. Once we form the hypothesis that the missing ingredient is causation, we can begin to recognize further counterexamples to the DN account. For instance, suppose that low atmospheric pressure causes both storms and high barometer readings (for simplicity, let's suppose that the causation here is deterministic).

    (a) Then, there will be a true lawlike universal generalization stating that, whenever the barometer reading is low, there will be a storm. So, we may form the following DN explanation of the storm:

    The barometer reading was high

    Whenever the barometer reading is high, there is a storm.

    There is a storm

(b) Even so, the low barometer reading doesn't explain *why* there was a storm. What explains *why* there was a storm is the low atmospheric pressure.

(c) This accords with our hypothesis—for the low atmospheric pressure *caused* the storm, and the barometer reading didn't.

13. We could, then, supplement the DN account with this new requirement to get the following theory of when a scientific explanation is good:

    The DN Account + Causation A good scientific explanation is one in which:

    (a) The explanans deductively entail the explanandum
    (b) The explanans are all true
    (c) The explanans have empirical content
    (d) The explanans contain at least one universal generalization which is a law
    (e) Without any given law statement, the remaining explanans do not deductively entail the explanandum
    (f) The explanans cause the explanandum

14. But then again, perhaps, once we have causation, we don't need anything else. Perhaps the other details of the DN account are superfluous once we have the causal condition.

15. In particular, Michael Scriven claims that merely causal information can provide an adequate scientific explanation, *without the need of any laws of nature*. For instance, he thinks the following explanation is a fine one:

    (a) Why is there an ink stain on the carpet? Because dopey Dr. Jones bumped his desk and knocked an ink bottle onto the carpet.

    But this explanation mentions only a *cause* of the ink stain. It does not additionally mention any laws of nature—nor is there any true law of nature to the effect that *whenever* Dr. Jones bumps his desk, there will be an ink stain.

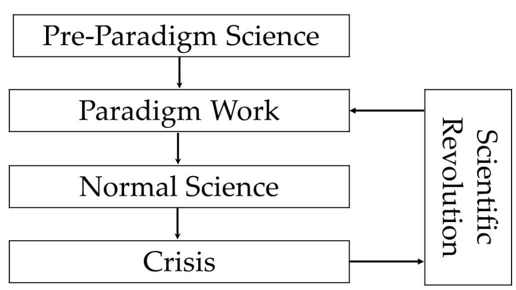
Part VI

Paradigms and Scientific Revolutions

23 Kuhn's Theory of Scientific Development

1. This week we will be discussing the historian of science Thomas Kuhn's book *The Structure of Scientific Revolutions*. In this seminal work, Kuhn is interested in many of the kinds of questions we have been considering in this course—what science can teach us about the world, how science justifies its theories, and so on. However, he is first and foremost interested in understanding the historical development of science.

2. In *Structure*, Kuhn presents a theory of the development of science. On this view, there are 5 important stages in the development of science: a pre-paradigm stage, the emergence of a paradigm work, a period of normal science, a crisis, and a scientific revolution in which the old paradigm is upended and a new piece of paradigm work emerges. We can visualize these stages of scientific activity in the following flowchart:



3. To understand Kuhn's theory, let's proceed through these stages of scientific development.

23.1 Pre-Paradigm Science

4. During the stage of pre-paradigm science, there are many competing schools of thought about how to understand some phenomenon of interest, with no clear con-

sensus about

- metaphysics (what there is);
- methodology (how to learn things about what there is);
- which questions are legitimate questions to ask; or
- how such questions are to be answered.

During pre-paradigm science, there is an intense focus on foundational questions.

5. For instance, prior to the publication of Newton's *Opticks*, there were at least two competing schools of thought about light and vision.

   (a) The *Eidolon* theory of light said that light was a tiny corpuscle (or particle) which objects *emit*, and which resemble that omitting object. These emitted corpuscles are then absorbed by the eye.

   (b) The *Emission* theory of vision, on the other hand, said that the eye emitted beams which could return information about external objects. Light makes the air permeable to these eye beams.

   (c) These schools competed with one another for a time and traded objections—but no clear consensus emerged about which theory was correct.

23.2 Paradigm Work

6. What Kuhn calls a 'paradigm' piece of scientific work is a scientific theory or explanation which clearly does better than its competitors.

   (a) This work will not in general be a complete and total theory of the target phenomenon. It will be partial and in need of completion/supplementation.

   (b) However, a paradigm piece of work will be work which *shows promise*, and clearly shows *more* promise than any other work being done at the time.

   (c) Once scientists appreciate this paradigm work, they take it to *set the agenda* for future work in the area. That is: they organize their intellectual activity around further developing and articulating the paradigm work.

7. For instance:

   (a) Newton's *Opticks* provided a theory of light as *rays*, along with a theory of the *reflection* and *refraction* of these rays. He used this theory to explain the colors of light in terms of the degrees of refection. With this theory in hand, he is able to give a compelling explanation of rainbows.

   (b) There were still many things that this theory didn't explain, but it clearly outdid its competitors. And it quickly *set the agenda* for future work in the area.

   (c) Thus, Newton's *Opticks* was a paradigm piece of scientific work.

8. For another example: Newton's *Philosophiae Naturalis Principia Mathematica*, which laid out Newton's laws of motion and the inverse square force law for gravitation, and used them to explain Kepler's laws about the speed of planets in their orbits.

9. A paradigm piece of scientific work will settle some answers to foundational questions which were unsettled during pre-paradigm science.

23.3 Normal Science

10. *Normal Science* is the stage in scientific development in which the community of scientists have organized their intellectual labors around the paradigm piece of scientific work. During normal science, scientists devote themselves to *articulating the paradigm*.

    (a) A word on Kuhn's terminology: his use of 'paradigm' is not consistent. We should distinguish between what we can call *paradigm in the narrow sense* and *paradigm in the broad sense*.

    (b) In the narrow sense, "paradigm" refers to the agenda-setting piece of scientific work.

    (c) In the broad sense, "paradigm" refers to the research program which is inspired by the paradigm piece of scientific work.

11. During normal science, there is a rigid consensus on:

    – Metaphysical questions about what there is (*e.g.*, there are light rays, or there are objects with masses, and those objects exert forces upon each other);
    – Methodological questions about how to determine what there is;
    – Which questions are legitimate; and
    – How those questions are to be answered.

12. The business of normal science is, according to Kuhn, an extended exercise in "puzzle solving". During normal science, Kuhn says that scientists:

    > "attempt to force nature into the preformed and relatively inflexible box that the paradigm supplies...those that will not fit are often not seen at all...Nor do scientists normally aim to invent new theories, and they are often intolerant of those invented by others."

13. During normal science, experimental work consists primarily of three kinds:

    (a) Experiments determining important values—*e.g.*, experiments which are used to calculate the value of G in Newton's force law, according to which the gravitational force which object 1 exerts on object 2 is given by

$$G \cdot \frac{m_1 \cdot m_2}{\vec{r}_{12}^2}$$

(b) Novel experimental confirmations of the existing paradigm; and

(c) Experiments meant to decide between different ways of articulating the existing paradigm—*e.g.*, experiments to determine the force law for electricity.

14. This rigidity of thought and intense focus on puzzle-solving is not a *vice*, in Kuhn's view. It is rather a *virtue* which is responsible for the monumental achievements of science. Kuhn writes:

> By focusing attention upon a small range of relatively esoteric problems, the paradigm forces scientists to investigate some part of nature in a detail and depth that would otherwise be unimaginable.

15. The intense focus on puzzle-solving means that the paradigm itself is largely shielded from refutation. The goal of normal science is to *develop* and *articulate* the paradigm—evidence which seems to disconfirm the paradigm is therefore taken as evidence against some auxiliary assumption made for the purposes of developing the paradigm, and not a refutation of the paradigm itself.

    (a) *Cf.* the Duhem-Quine Thesis.

    (b) Thus, while smaller, subsidiary hypotheses can be empirically refuted within the paradigm, the paradigm itself enjoys a presumption of innocence. (*Cf.* Popper's refutationism)

16. Sometimes, a paradigm runs into trouble—there are "anomalies": phenomena which the paradigm does not seem to be able to explain, or puzzles the normal scientists are not able to solve.

    (a) For instance, Newtonian mechanics was not able to fully explain the precession of the perihelion of Mercury.

17. As long as these anomalies are few, normal scientists do not take them as reasons to reject the paradigm.

18. When, however, a critical mass of anomalies amasses, scientists begin to lose their faith in the paradigm, and science is plunged into a period of *crisis*.

    (a) For instance, towards the end of the nineteenth century, physicists were *very* confident of the paradigm of Newtonian mechanics which they had been articulating since the publication of Newton's *Principia* nearly two centuries prior. They had concluded that light was an electromagnetic wave—but a wave must propagate through some medium, which they called the *luminiferous aether*. So Michelson and Morely set out to measure the earth's motion through this luminiferous aether. When they did so, they could not detect any motion. It seemed that light was moving at the same speed in every direction.

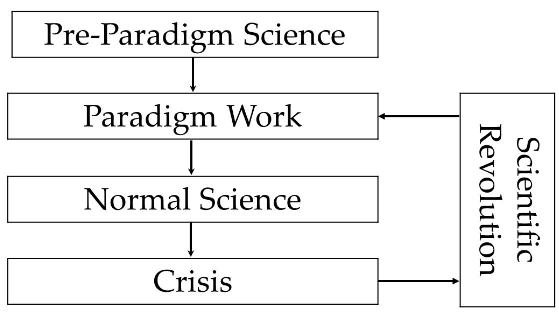
    (b) The outcome of the Michelson-Morely experiment was an upsetting anomaly which began to shake physicists' faith in the prevailing paradigm of Newtonian mechanics.

23.4 Crisis & Scientific Revolution

19. During this period of crisis, scientists once again take an interest in foundational questions about
    - Metaphysics (what there is);
    - Methodology (how we should determine what there is);
    - Which questions are legitimately asked; and
    - How such questions are to be answered.
20. Nevertheless, even during the stage of crisis, scientists will continue to cling to the paradigm theory. They will not abandon the paradigm until the discovery of a new theory which is capable of improving upon the previous paradigm—and, in particular, dealing with the anomalies turned up in the old paradigm.
    (a) E.g., Einstein's theory of relativity replaces the Newtonian paradigm put in place by the *Principia*.
21. When a new paradigm (in the *narrow sense*) emerges, there is a *scientific revolution*. Scientists abandon the old paradigm and move to the new one. The new paradigm once again settles the foundational questions about what exists, how it is to be investigated, which questions are legitimate, and so on, and normal science takes over again under the aegis of the new paradigm.

24 | Kuhn on Scientific Revolutions

1. Recall, Kuhn's theory of the development of science involves five stages: first, there is a stage of pre-paradigm science; once a paradigm work emerges, scientists begin to organize their work around the articulation of the paradigm—this stage is what Kuhn calls *normal science*. Once a critical mass of anomalies begins to amass, normal science gives way to a period of crisis; finally, once a new paradigm work emerges, there is a scientific revolution, and normal science begins anew.



2. At first glance, Kuhn's views about the development of science are perfectly compatible with saying that science *progresses*, and not merely that it *develops*.
3. However, in his final chapters, Kuhn raises some reasons to doubt that science progresses in any meaningful sense.
   (a) The main problem for Kuhn was that, in order to say that science has *progressed*, we must say that the new paradigm is better than the former paradigm. However, Kuhn thought that the means by which we might evaluate the old and the new paradigms is itself *paradigm dependent*.
   (b) For this reason, Kuhn thought that different paradigms are *incommensurable*—that is, that they are incapable of being compared.

24.1 Incommensurable Meaning

4. Kuhn accepted a *holistic* theory of the meaning of theoretical terms. That is, Kuhn thought that the only way to determine the meaning of a theoretical term like *electron* or *mass* was to look to the entire theory in which the term appears.

   (a) That is, the *meaning* of "mass" is determined in part by the theories speakers accept.

   (b) Since Newtonian mechanics and Relativity treat "mass" very differently, Kuhn concludes that "mass" *means something different* inn Relativity than it does in Newtonian mechanics.

   (c) Call this Kuhn's thesis of *Semantic Incommensurability*.

5. Why does the thesis of semantic incommensurability call into question that science makes progress? Consider the claims made by Newtonian mechanics and Relativity:

   Newtonian Mechanics: "An object's mass does not depend upon its velocity".

   Relativity: "An object's mass does depend upon its velocity".

   Now, we might want to say that science has progressed because Newtonian mechanics accepted a *false* claim. However, we cannot infer that the claim made by Newtonian mechanics is false simply because the claim made by Relativity is true—for, in the claim made by Newtonian Mechanics, the word "mass" *means something different* than it does in the claim made by Relativity. Compare:

   Alice: "My wife is at the store"

   Bob: "My wife is not at the store"

   Alice and Bob are not contradicting each other. For, when Alice speaks "my wife" *means something different* than it means when Bob speaks. So we can't infer that Alice has said something false just because Bob says something true.

24.2 Incommensurable Standards

6. The thesis of Semantic Incommensurability poses a problem for our ability to say that the claims of rejected theories have been shown to be *false*. But perhaps we can still say that we are doing *better science* in the new paradigm than we were doing in the old paradigm.

7. Here, too, Kuhn sees trouble. For Kuhn believes that each paradigm comes with its own standards for what constitutes *good* science.

(a) For one example, prior to Newton's theory of gravitation, it was widely held that a good scientific theory would provide a *mechanistic* explanation of phenomena. However, Newton's theory lacked any mechanistic theory of gravity. It appeared that gravity operated instantaneously at a distance. In the pre-Newtonian era, this would have been viewed as bad science. However, from the standpoint of the new paradigm, it was viewed as good science.

8. Call the thesis that each paradigm has its own standards for what constitutes good science the thesis of the *Incommensurability of Scientific Standards*.

9. Kuhn takes the incommensurability of scientific standards to mean that, when we say something like "Quantum Mechanical explanations of the behavior of light are better than the explanations of Newton's ray theory or Huygen's wave theory", we are, to some degree, presupposing the standards of our own paradigm.

10. Note that we could understand Kuhn's thesis of the incommensurability of scientific standards in two different ways.

    (a) On the one hand, we could take him to be merely describing the attitudes of the scientists. That is, the view could be that the standards with which scientists evaluate theories as good changes from paradigm to paradigm.

    (b) On the other hand, we could take him to be describing the true standards. That is, the view could be that the standards which *genuinely make science good* change from paradigm to paradigm.

11. Nothing that Kuhn has said about the history of science gives us a reason to conclude that the actual standards have changed, rather than the scientist's attitudes.

    (a) Moreover, if the attitudes of the scientists are getting closer to the actual standards, then surely this constitutes some form of progress.

24.3 Incommensurable Observations

(a) Semantic incommensurability is meant to pose difficulty for our ability to say that the claims of superseded paradigms were false.

(b) The incommensurability of standards is meant to pose difficulty for our ability to say that superseding paradigms are *better science* than superseded paradigms.

12. But at least we should be able to say that superseding paradigms are *better supported by the evidence*—right?

13. Kuhn is not so sure. He thinks that the content of our observations—and therefore the evidence for which a theory has to account—will also depend upon the theories we accept.

    (a) For instance, what we make of the evidence provided by a microscope will depend upon whether we accept theories of light which tell us that the microscope accurately presents microscopic reality, or whether we accept theories which tell us that they distort microscopic reality. And similarly for our observations of X-rays, trails in a cloud chamber, and so on. If we accept the current paradigm, we will see a tumor, or positrons. If, however, we accept a different theory, then we may see something different.
    (b) *Cf.* the Necker Cube or the Duck-rabbit.
    (c) Call this the thesis of the *theory-dependence of observation*.

14. Thus, Kuhn thinks that 1) the meaning of our theoretical terms; 2) the standards by which we judge science; and 3) our observations are each infected by our current paradigm.
    (a) Semantic incommensurability means that whether the previous paradigms were *false* is a paradigm-dependent matter;
    (b) incommensurability of scientific standards means that whether superseding paradigms are *better science* is itself a paradigm-dependent matter; and
    (c) the theory-dependence of observation means that whether superseding theories accord better with the evidence is itself a paradigm-dependent matter.

24.4 Off the Deep End

15. In the infamous chapter X of Kuhn's *Structure of Scientific Revolutions*, he makes the following bold claim:

    > *In so far as their only recourse to [the] world is through what they see and do, we may want to say that after a revolution scientists are responding to a different world*

    and, again:

    > *The very ease and rapidity with which astronomers saw new things when looking at old objects with old instruments may make us wish to say that, after Copernicus, astronomers lived in a different world.*

16. We might try to read these passages metaphorically—and this is how Kuhn said they should be read later on—but passages like the following suggest the more radical reading:

    > *Many readers will surely want to say that what changes with a paradigm is only the scientist's interpretation of observations that themselves are fixed once and for all by the nature of the environment and of the perceptual apparatus.*

After this, Kuhn suggests that this urge is itself a feature of our prevailing paradigm—however, one that may be in the process of shifting; as he says,

> *Today, research in parts of philosophy, psychology, linguistics, and even art history, all converge to suggest that the traditional paradigm is askew*

17. Despite Kuhn's disavowals of the radical reading of these passages, they have been wildly influential. For instance, here is the sociologist Bruno Latour, claiming that Louis Pastour did not *discover* microorganisms, but rather *invented* them:

> *'But', anyone with common sense would ask with an undertone of exasperation, 'did ferments exist before Pasteur made them up?' There is no avoiding the answer: 'No, they did not exist before he came along'.*