CHAPTER 1: THE MEANING OF LIFE

A. Life’s Chronic Ailments
   - Gilgamesh and Death
   - Sisyphus and Futility
   - Job and Suffering
B. Ancient Greek Solutions
   - Epicureanism and Pleasure
   - Stoicism and Accepting Fate
   - Skepticism and Doubt
   - Cynicism and Defying Convention
C. Western Religious Solutions
   - Having Children
   - Life after Death
   - Furthering God’s Kingdom
D. Eastern Religious Solutions
   - Daoism and the Way of Nature
   - Buddhism and Extinguishing Desire
   - Hinduism and the Four Goals of Life
E. Philosophy and Life’s Meaning

For Reflection

1. What kinds of things might cast a shadow over the meaningfulness of life?
2. Is the meaning of life for a human substantially different from the meaning of life for a cat?
3. Suppose that God doesn’t exist and human beings are simply the result of accidental evolutionary forces. Does this make life meaningless?
4. What are some religious beliefs or pursuits that might give meaning to life?
5. What are some non-religious beliefs or pursuits that might give meaning to life?

6. Compare a typical city-dweller today to a typical hunter-gatherer 20,000 years ago. Which of the two would have an easier time finding meaning in life?

John is a fervent Star Wars fan and has devoted much of his adult life to attending Star Wars conventions—dressed as a storm trooper—and contributing to Star Wars websites. He keeps a particularly close eye on the main “Star Wars” entry in Wikipedia to weed out factual errors and cleanse it from what he calls “Light Side of the Force ideological bias.” John’s pride and joy is his collection of nearly 500 Star Wars action figures, including three design concept ones used in film production. It’s not merely a toy collection, he argues, but a way of exploring the meaning of life:

The Star Wars narrative contains all of the major plot motifs of classic literature, and the action figures give reality to them. How I place action figures together on the shelf will evoke different thematic tensions based on their respective back stories. When I pair Han Solo with Greedo, that displays a completely different Light-side dark-side dichotomy than when pairing Hans Solo with, say, General Grievous.

While some people find meaning through religion or acquiring wealth, John says that he find Star Wars action figures to be a much more flexible and organic expression of life’s relentless struggles and how we meet them. John’s wife, who he first met at a Star Wars convention, tolerates his collection of action figures, but thinks it’s a little too intense: “I feel like I’m surrounded by 500 tiny warriors poised for battle; I can think of more soothing decorations to place around the apartment.”
- Does life have a purpose?
- What kind of life is worth living?
- How can I overcome despair?
- How can I achieve happiness?
- Why do I exist?
- Why should I exist?
- Do my life activities have any lasting value?

Each of these questions focuses on a unique point. The first, for example, asks whether there is an overarching design or goal to human existence that might clarify my place in the grand scheme of things. The second asks whether some approaches to life are better than others. All of the above questions, though, presume that something’s not quite right with life as we currently experience it, and we’d like a solution to the problem.

Philosophers are not the only ones interested in questions about life’s meaning. Today especially there is a large self-help industry devoted to solving personal struggles. While many of these involve quite specific problems—relationship issues, alcohol dependence—others are more general in nature. A mid-life crisis or a “spiritual” crisis, for example, will often involve larger questions of purpose and fulfillment. Philosophical discussions of the meaning of life are not meant to compete with self-help therapies. The main appeal of philosophical approaches to life’s meaning rests in the puzzle itself: here is a timeless problem that touches the very core of human existence. What exactly is behind the problem and which, if any, of the stock solutions are plausible?

In this chapter we will look at the more famous problems and proposed solutions regarding life’s meaning that have attracted the interest of philosophers over the millennia. Many of the proposed solutions come from ancient traditions—both religious and non-religious. To get a complete picture of their approaches to life’s meaning, we’d need to immerse ourselves in all the particulars of those traditions and the precisely defined lifestyles that they advocate. The best we can do here, though, is consider some dominant themes of these traditions, along with some common criticisms of them. The criticisms we’ll look at are by no means refutations of those traditions, and advocates of those traditions would certainly have responses to them. Rather, the critiques serve more to help define their limits rather than to simply dismiss them.

A. Life’s Chronic Ailments

Preoccupied with our own private problems, it’s easy for us to forget that for about 100,000 years human beings just like us have been on this planet, undoubtedly wrestling with their own issues of happiness and contentment. It should be no surprise that, as soon as writing was invented, ancient people inscribed their struggles to find life’s meaning. Their accounts don’t typically begin “Today I had a really bad day”; instead, writers embedded their insights into mythological narratives, the popular writing genre of the time. Three ancient discussions are especially outstanding because of their insight and influence, and each describes a particular obstacle that stands in the way of us having a meaningful life.

**Gilgamesh and Death.** One of the world’s oldest surviving stories is the Epic of Gilgamesh, composed about 4,000 years ago in Mesopotamia. Gilgamesh, a brave and heroic king, just witnessed the death of his close friend and became distressed with the prospect that he too would someday die. There’s got to be some cure for death, he thought, and so he set out on a journey to discover it. Everyone he encountered on his travels, even animals, tried to discourage him from pursuing his plan, but he pressed on all the same. He then found a famous man named Utnapishtim who had himself achieved immortality. Utnapishtim was the Mesopotamian Noah who survived the great flood. Warned by a goddess of the forthcoming deluge, Utnapishtim built a ship to save himself and his family; he was granted immortality as a reward for
his efforts. Gilgamesh was shocked when he first set his eyes on Utnapishtim, who, while immortal, continued to age. The old man was now so decrepit that he could barely move. Gilgamesh nevertheless asked his advice and Utnapishtim offered a suggestion: Gilgamesh could conquer death by staying awake for seven nights straight. Gilgamesh accepted the challenge, but fell asleep as soon as he sat down. When awakened, he was prepared to return home without success. Utnapishtim’s wife then urged the old man to tell Gilgamesh about a secret cure for death: at the bottom of the ocean there’s a spiky plant that brings endless youth to anyone who eats it. Gilgamesh ran to the ocean, and, with rocks tied to his feet, jumped in and sank to the bottom. He grabbed the spiky plant, untied the rocks and floated back to the surface. Plant in hand, he joyfully set out on his return journey and when almost home he stopped to wash himself off in a stream, first placing the plant on the bank. While bathing, though, an old snake slithered up to the plant, ate it, and immediately became young; it then slithered away. Gilgamesh’s one chance at becoming immortal was thus ruined. He arrived home in a state of depression, and, in spite of the efforts of his friends to cheer him up, he remained inconsolable.

There are two morals of this story. The first and obvious one is that, as strongly as we desire to live forever, the inevitable truth is that we will all die. Given the choice, virtually all of us would jump at the chance to live forever, and the fact that we can’t creates a dark cloud over life’s meaning. The second and more interesting moral is that we cannot easily accept our deaths and we may do some crazy things to cheat the grim reaper. While the epic of Gilgamesh is just a myth, this second moral has played out countless times in the real world. In ancient China, many religious believers devoted themselves to conquering death through the strangest of techniques. One involved drinking chemical concoctions which would supposedly balance out the forces within the human body and thereby obstruct the process of dying. Ironically, many believers poisoned themselves to death through these experiments. Another technique involved holding one’s breath for longer and longer periods of time. Eventually the believer would not need to breathe at all, and thereby become immortal. Today, several organizations are devoted to achieving physical immortality. Some recommend taking as many as 250 nutritional supplements a day. Others place hope in biological advances that will reverse the natural deterioration of human cells. Still others look forward to the day when our minds can become digitized—essentially making computerized versions of our present brain processes.

What should we think about these efforts to avoid dying? One of the more notable philosophical discussions of death is by German philosopher Martin Heidegger (1889-1976). Death, according to Heidegger, is not really an event that happens to me, since it just involves the termination of all possible experiences that I might have. After all, it is impossible for me to experience my own death. Rather than thinking of death as an episode that takes place at the tail end of my life, I should instead view it as an integral part of who I am right now, and during each moment of my life in the future. I continually aim towards death and, even when I feel healthy, in a fundamental way I’m really terminally ill. It’s like playing a game such as soccer where, embedded in every moment, there is the idea that time is running out. So, Heidegger says, if I ignore my persistent movement towards death—or resist it as Gilgamesh did—I’m just deceiving myself and living in a substandard world of make-believe. By contrast, a proper understanding of death clearly lays down the basic rules of the game of life and thereby gives life form and purpose.

If I could continually think of myself as on the path to death as Heidegger suggests, that might help me accept my mortality. However, while my body is clearly designed to die, my mind seems...
to be hardwired to think that I’m immortal, and there’s little that I
can do to resist that feeling. For one thing, the natural instinct to
survive compels me to resist death at almost all costs; this is
something that I share with many creatures in the animal world.
For another, I can’t psychologically conceive of the future without
secretly injecting myself into it. Even if I try to picture the world a
thousand years down the road, I’m still there as a kind of ghostly
spectator to the events I’m imagining. Whether I like it or not, I’m
inherently resistant to the idea of my non-existence. My natural
human attitude towards death, then, may be to simply assume that
I’m immortal, and, at the same time, be horrified when I look down
and see my body disintegrate before my eyes. So, anxiety over
death, like Gilgamesh experienced, may just be part of life.

_Sisyphus and Futility_. In Homer’s _Odyssey_, the adventurous hero
Odysseus stops by Hades—the dwelling place of the dead—to chat
with deceased friends. While there, he sees several legendary
people who are being punished for evils they committed when
alive. There’s one fellow whose body is stretched out over a nine
acre area. Lying helplessly, two vultures pick at his liver; he swats
them to shoo them away, but they keep returning. Another fellow is
parched with thirst, but cannot succeed in reaching water. Wading
in a lake up to his chin, whenever he stoops down to drink, it
immediately dries up leaving only dusty ground. He sees succulent
fruit trees above him, but as soon as he reaches for their produce
the wind sweeps the branches into the clouds.

And then there is Sisyphus, a deceitful king who tricked the god of
death and stayed alive longer than he should have. He finally died
and went to Hades, but the punishment for his trickery was not a
pleasant one. Day after day he pushes a huge stone up a hill, but,
always losing energy as he nears the top, he lets it go and it rolls
back down. Homer describes the scene here:

```
I saw Sisyphus at his endless task raising his gigantic stone with
both his hands. With hands and feet he tried to roll it up to the
top of the hill, but always, just before he could roll it over onto
the other side, its weight would be too much for him, and,
without pity, the stone would come thundering down again onto
the plain below. Then he would begin trying to push it up hill
again, and, as the sweat ran off him, a steam rose after him.
[Odyssey, Book 11]
```

All of these scenes from Hades depict people
trapped into performing futile tasks—swatting
vultures, stooping to drink, pushing a boulder.
It’s the image of Sisyphus, though, that’s had
the most lasting impact, and for nearly 3,000
years writers have used him as a symbol for the
emptiness of life’s endeavors.

Sisyphus’s fate is frighteningly similar to the assembly line jobs
that workers face throughout the world. Jill works in a lawnmower
manufacturing plant, and her job is to bolt lawnmower blades onto
motors. She has thirty seconds to line up the pieces and attach them
together. As soon as one is done, another follows on its heels. To
reduce monotony, the factory rotates Jill and other employees from
one work station to another, but, after a few minutes, the routine
kicks in. Jill likes her co-workers and has no complaints against her
supervisor. Still, at the end of the day, she feels that she may as well
have been pushing a boulder up a hill. It’s not just assembly line
jobs that carry a sense of tedious futility. Accountants, teachers,
doctors, and virtually all skilled workers face early burnout.
Surveys show that about 70% of workers dislike or downright hate
their jobs—much of which owes to grinding and pointless routines.
What we do in our spare time is often no more rewarding. A good
portion of the day is spent in monotonous domestic chores—
cleaning, driving to and fro, shopping—which, year after year,
seem as futile as assembling lawnmower blades.
French philosopher **Albert Camus** (1913-1960) believed that the story of Sisyphus had another symbolic message. Yes, many of life’s specific tasks certainly feel futile; however, what’s more discouraging is that the sum total of a person’s life efforts may seem pointless. Camus called this *the absurdity of life*. Human life, he argued, cannot be neatly dissected and understood by human reason in the same way that scientists might successfully analyze and understand chemical reactions. We strive to be happy, but instead are trapped in a life of futile efforts. As much as we try to make sense of it and solve the problem, we can’t; the sober reality of things simply does not live up to our optimistic expectations. The problem is so bad that it might drive some to suicide. So, Sisyphus represents the overwhelming struggle that we each have in overcoming a pointless life. But Camus is not content to let the issue rest with despair. Instead, he recommends that we revolt against the apparent pointlessness of life, accept our condition as limited as it is, and in that find happiness. Sisyphus should embrace his boulder-pushing endeavor; the value rests in his effort, not in what he achieves. Similarly, while I may never be able to rationally explain the purpose behind my life, I should nonetheless welcome the life that I have and create meaning for myself through my positive outlook.

But is Camus’ recommendation as easily achievable as he thinks? That is, through sheer willpower can we really make ourselves happy in spite of life’s fruitlessness? The problem may be resistant to a simple attitude adjustment, as zoo keepers have recently discovered in their experience with the mental well-being of gorillas. For decades gorillas were kept in controlled enclosures with fixed routines like feeding schedules. While their basic needs were being met, the gorillas were all bored and depressed. Zoologists then discovered that gorillas needed sufficiently complex tasks to challenge them throughout the day and keep their mental energies peaked. Caretakers then started regularly altering the gorillas’ routines by introducing different climbing equipment and scattering their food around their enclosures for the gorillas to forage. By altering the gorillas’ environment in the right way, the gorillas became happier. Applying this lesson to human happiness, we might look for the kinds of challenging tasks that spark our interests throughout the day. We might need shorter and more varied work days; we might need more direct involvement with growing and preparing food; we might need the opportunity to explore new surroundings through travel; we might need to break free of overcrowded urban settings. In the end we might find that humans were designed to be content in tiny hunter-gatherer tribal groups—the condition in which the human species first evolved. In any event, modern industrial life may just not be suited to ward off a sense of futility, and for us the human condition today may be inherently absurd with no real solution. Like Sisyphus, then, we unendingly push a boulder to no avail.

**Job and Suffering.** The story of Job from the Hebrew Old Testament explores another challenge to the meaning of life. Job was not obsessed with death like Gilgamesh or disheartened from futility like Sisyphus. In fact, at the outset of the story he’s quite happy. Job is a wealthy and morally decent herdsman with a loving family, and he owns a large stock of sheep, oxen, camels, and donkeys. Then everything changes for the worse. His animals are stolen, his servants are burnt to death by fire from the sky and, worst of all, his children are killed in a tornado. Job himself is infected with itchy skin boils, which he scratches with a broken piece of pottery. In a display of sorrow, he rips his clothes and shaves his head. Three friends stop by for a visit and at first don’t even recognize Job because he’s so disfigured from his illness. For a week they sit next to him without speaking, then, breaking the silence, Job says “I wish I was born dead!” He can’t understand why God would do this to him, and he accuses God of being his tormenter. His friends try to explain God’s role in his misfortunes. One friend argues that people suffer when they forget God and, so,
Job must have abandoned God at some point in his life. Another argues that people suffer when they commit some moral offense, and no one can fully know all the things that God finds evil. So, in spite of Job’s protests of being morally blameless, he nevertheless must have committed some offense that isn’t immediately apparent. Job insists, though, that he did nothing wrong. Finally, God himself appears in a thunderstorm and sets the record straight: God is infinitely great, Job is virtually insignificant and, so, Job has no right to complain.

The problem raised in the story of Job is how we explain human suffering. While all suffering is inherently bad, though, it’s only a specific type of misery that casts a serious shadow over the meaning of life. Suppose I pick up a hammer and intentionally hit myself on the foot with it. The explanation of my suffering is clear and there is no moral mystery to be solved: I have no one to blame but my foolish self. This is a rule of life that I understand and accept, no matter how miserable I make myself. Suffering of this sort, then, poses no real threat to a meaningful life. It may not even be so bad if you intentionally hammer away at my foot, so long as you are arrested and convicted of assault. Even though I’m in pain, I can be consoled by the fact that justice has been done and you are held accountable for my suffering. So, even unjustified suffering like this won’t necessarily make my life meaningless. The real problem occurs when the suffering is both unprovoked and unresolved, which is exactly what Job faced. In spite of his friends’ accusations, Job was convinced that he did nothing to deserve his suffering; from his perspective, it was completely unprovoked. It was also unresolved since, when his livestock was stolen the bad guys got away with it. If they had been arrested and forced to compensate Job for his losses, then perhaps Job could have accepted the situation and moved on. Job was not so lucky, though. Similarly, when his children were killed, he could not just replace his old family with a new one. He also could receive no consolation prize that would compensate him for his agonizing illness.

With no resolution to these unprovoked tragedies, Job is left wondering why they happened. Part of human nature is to seek out the hidden causes of things and resolve mysteries. When tragedy strikes us through no fault of our own, we are inclined to find some cause and, more importantly, cast blame on that cause when we can. This is one reason why lawsuits are so common today. If I trip over a curb, it’s the city’s fault for placing it where they did, and I sue them. If I fall off a ladder, it’s the ladder company’s fault for not warning me about possible dangers, and I sue them. If Job had the chance, he might have sued his local police for not catching the thieves, or sued the National Weather Service for not forewarning him of the tornado. But the more irrational our accusations are, the less comfort we can take in them, and, in our more clear-headed moments, we are still left wondering why these tragedies happened. When we fail in our attempts to find blame with human causes for our misery, many people, like Job, cast blame on divine causes. An all powerful God should protect me from unprovoked suffering, and if he doesn’t, then he’s to blame. In Job’s words, God is our tormentor.

Job’s story ends with a resolution of a sort. He directly witnesses God’s vastness and then grasps the enormous gulf between the two of them; the very experience of divine power humbles him to accept his situation. But this is a storybook ending, since most believers will not have a direct experience of God’s greatness to force them in line. Imagine that you lost a relative in a tornado and you put the blame on God. In my efforts to comfort you I said, “I know this is painful for you, but don’t be discouraged. God is infinitely great and you are by comparison insignificant; this is what we learn from the story of Job.” This would offend you more than it would console you. In the course of our lives, most of us experience tragedies that are unprovoked and unresolved—property loss, the
death of loved ones, serious illness. We can surely appreciate Job’s despair when no satisfactory explanation is available.

B. Ancient Greek Solutions

Just as problems with the meaning of life were voiced early on in human civilization, so too did the ancient world propose solutions. The theories that they suggested were quite varied, and we’d be hard pressed to find a solution today which wasn’t first entertained back then. The first set of solutions we’ll look at are from ancient Greece. For a brief period of time, Greek philosophers were in the self-help business and they offered step-by-step methods for achieving happiness. Four approaches were so popular that even today their names are household words: Epicureanism, Stoicism, Skepticism, and Cynicism. Here are some of their main themes.

Epicureanism and Pleasure. Jack, an English professor from a large and prestigious university, thinks he’s cracked the code to happiness. Divorced and in his mid 40’s, he makes a good income at a job that doesn’t require much work. He published a lot earlier in his career, but now he rides on his reputation and gets by doing minimal preparation for the few classes that he’s required to teach. In his spare time he indulges his many cravings. An enthusiast of specialty foods, he is intimately familiar with the menus of every fine restaurant in his area and he regularly attends wine and cheese tasting events. During the day he reads novels, plays tennis, visits art museums, and takes sculpting classes. In the evening he watches foreign films at art houses, after which he frequents local jazz clubs. On school breaks he flies to Europe, sampling the cultural offerings there. His passions, though, are not limited to food, art and travel. Jack possesses an animal magnetism that makes him particularly successful in the romance department. Each semester he invites a new female graduate assistant to be his lover for the duration of the term. While the women know that the affair is only temporary, they happily agree, and even recommend possible partners for his next semester. On his birthday, his ex-lovers who are still in the area throw him a party. In a word, Jack is an Epicurean.

The Greek philosopher Epicurus (341-271 BCE) believed that the job of philosophy is to help people attain happiness; a philosophy that does not heal the soul, he argues, is no better than medicine that cannot cure the body. His formula for attaining human happiness is simple: increase pleasure and decrease pain. Personal pleasure is the only thing that we should pursue, and the value of everything we do in life is judged by that standard. The pleasures that Epicurus recommends are precisely the ones that Jack enjoys, but he warns that we should not pursue all pleasures with equal zeal. First, some are physical such as Jack’s romances, and others are mental such as Jack’s love of art; the mental ones are more important than the physical ones. Second, some desires are not entirely necessary, such as the desire for luxury food, and we should pursue these with moderation. Third, Epicurus warns us to avoid placing short term desires above long-term ones. For example, if Jack skipped teaching his classes for the short term goal of visiting a museum, then he’d likely lose his job and his happy lifestyle would come crashing down.

Is Epicureanism a reasonable path to human happiness? While we all naturally want pleasure, there is something suspect about a lifestyle that is devoted entirely to its pursuit. Let’s grant that Jack is truly happy with his Epicurean existence. There’s no telling, though, how long those activities will sustain his interest. Part of the joy he experiences comes from the newness of his activities: a new restaurant, a new art exhibit, a new story plot, a new lover. There are only a finite number of spices to mix into one’s food, though, and eventually even the most unique of Jack’s experiences will take on familiar patterns and become routine. He’ll be like Sisyphus pushing a gem-encrusted boulder up a hill, a task no less
futile than pushing an ordinary rock. Further, the happiness that Jack does experience rests on a stroke of good fortune that may easily change. If his university cracks down on his laziness, he’ll have far less leisure time for his hobbies. If his ex-wife sues him for alimony, he won’t be able to cover the costs of his activities. As he grows older, young women will be repulsed by his romantic advances. Thus, indulging in pleasure is not a very stable road to happiness if it rests on so many factors beyond our control.

Epicurus himself was very restrained in the pleasures that he pursued. He lived on a small food diet, avoided luxuries, and strived for self-sufficiency. “The greatest benefit of self-sufficiency,” he argued, “is freedom.” It seems, then, that the founder of this pleasure-indulging lifestyle was far less Epicurean than we might think, and, instead, he grounded his happiness upon a feeling of independence. Thus, pursuing pleasure alone is no guarantee of a meaningful life, which Epicurus himself recognized.

Stoicism and Accepting Fate. Imagine that you are a captured soldier detained in a prisoner of war camp. Not particularly fond of the Geneva Convention, your captors have provided you with stark and sometimes inhumane accommodations. Your cell block is unheated, your bedding is covered with fleas, your meals are unpredictable and, when they are served, the food is often rotten. About once a week you are interrogated by your captors, who psychologically intimidate you and sometimes beat you. You don’t know how long your detention will last, or even if you’ll survive. In these conditions, could you possibly be happy? For starters, you’d have to condition yourself to ignore the physical harshness of your environment. Mustering all your mental strength, you might eventually get used to your cold room, unsanitary bedding and disgusting food. You’d then have to accept that you are at the mercy of the unpredictable whims of your captors who can beat you and even kill you as they see fit. Having no expectations at all about circumstances beyond your control, you might eventually be able to carve out some peace of mind. This is precisely the Stoic philosophy for achieving happiness. While life is not always as grizzly as a prisoner of war camp, sometimes it really is that bad, and there’s nothing we can do about it. If we place our hopes in pleasures that are beyond our control, we will inevitably be frustrated and unhappy. The moral of the story is that we should learn to accept the life that is fated for us, and never reach beyond that.

One of the great teachers of Stoicism was Epictetus (c.55—c.135 C.E.), a former slave who knew first hand how brutal and unpredictable life could be. He offers an especially picturesque example to explain the Stoic solution. Think of life as a large banquet with many people sitting around a table waiting to be fed. Starting at one end of the table, serving dishes of food are passed around, and guests scoop out portions onto their plates. You are near the end of the table and for all you know the serving dishes will be empty by the time they reach you. You should not keep glancing down the table in anticipation, Epictetus advises, but wait patiently for your turn. Better yet, he says, when a serving dish finally arrives, you should just pass it along without taking anything. This is what our attitudes should be toward the things in life that we typically crave—good jobs, a loving family, luxuries—but which we can never count on. For this Stoic formula to succeed, we must learn to habitually distance ourselves from things that we desire, even when things are going our way. The goal is to acquire a constant mental state of detachment so that, in the event that circumstances sour, we won’t be disappointed.

The Stoic path to happiness seems well suited for prisoners of war, slaves, and the financially destitute. For these people, life’s prospects are so dismal that placing hope beyond themselves will only make matters worse. What about the rest of us, though, who can at least occasionally count on good fortune at life’s banquet
Sometimes I will indeed be disappointed when a serving dish comes around empty, but this may well be counterbalanced by joys I will experience when another serving dish is full. For example, when hunting for a job, I will undoubtedly be disappointed if a company rejects my application, but I can reasonably expect that some company will hire me, and it doesn’t hurt to anticipate that with hope. Epictetus thought that one of life’s biggest disappointments was the death of a loved one. His Stoic recommendation is that we should emotionally distance ourselves from our spouses and children so that, when fate unpredictably tears them away from us, we won’t be distressed. Here again, though, while the death of loved ones is devastating, it is nevertheless counterbalanced by the joy we receive from our attachment to them while they are alive. And this is an important joy in life that we would sacrifice if we followed his Stoic advice. Stoicism, then, seems to be an unnecessarily extreme and restricting avenue towards happiness, which we should adopt only as a last resort when things become overwhelmingly dismal.

Skepticism and Doubt. An organization called “The Skeptics Society” is devoted to debunking a host of questionable beliefs, including UFOs, alien abductions, extrasensory perception, religious miracles, time travel, and conspiracy theories. One writer for the society cast his skeptical eye on the famed 1947 alien space craft sighting in Roswell, New Mexico. The real event, he explains, was simply a military balloon experiment, which decades later was transformed into a UFO legend. He writes,

Roswell is the world’s most famous, most exhaustively investigated, and most thoroughly debunked UFO claim. It’s far past time for UFOlogists to admit it and move on. Those who hope to discover alien life are going to have to look where the aliens are—which is (if anywhere), somewhere else. Perhaps outer space would be a good place to start. [B.D. Gildenberg “Roswell Explained”]

By exposing the faults in controversial claims such as the Roswell incident, The Skeptics Society hopes to promote critical thinking and proper scientific inquiry. The Society sees itself as following in a long skeptical tradition that began in ancient Greece, particularly the school of Skepticism founded by the philosopher Pyrrho (c.365-c.275 BCE).

Pyrrho and his followers maintained that happiness is achieved through doubt. The sort of happiness that they envisioned was mental tranquility—a peace of mind that we experience when we suspend belief. When we hold extreme views, such as belief that aliens visited Roswell, we experience a mental disturbance, and we risk being pulled from one conviction to another. If the aliens did appear there, what was their mission? If the government knew about the event, why are they covering it up? We quickly become tangled in a web of questions and concerns. It’s not just strange beliefs like this that disrupt us, but any strong conviction upsets our peace of mind when we hold rigidly to it, such as the belief that the grass in my yard is green or that the table in my kitchen is round. The solution, according to the skeptics, is to recognize that every belief is subject to doubt. The grass appears green because my eyes are constructed a specific way and light shines on it in a specific way. If these factors differed, then the grass would not appear green. So, I should suspend belief about whether the grass is really green. Skeptics argue that I should in fact suspend all beliefs that I hold—about external objects, God’s existence, moral values—and thereby free my mind of the conflict that these beliefs produce. By achieving this mental tranquility, I will become happy.

The skeptic is probably right that the more gullible we are, the more we set ourselves up for disappointment. By believing in UFOs, horoscopes or miracle cures, we place ourselves at odds with respectable methods of inquiry and invite ridicule. If I persist in my strange beliefs, contrary to strong evidence against them, then I
must brainwash myself in thinking that I’m right and everyone else is wrong, which is a discomforting prospect. But the skeptic’s larger point is that all beliefs—both strange and normal—are vulnerable to attack and should thus be rejected for the advancement of mental tranquility.

There are two problems with this position. First, suppose that the skeptic is right that even our most commonsensical beliefs can be called into question, such as the belief that the table in front of me is round. It’s one thing for me to recognize the theoretical problems with that belief; it’s entirely another thing to actually suspend my belief about the table’s roundness, especially when it always appears to me that way. Commonsense beliefs like this may simply be beyond my control, regardless of how hard I try to suspend them. I’m forced to act on the assumption that the table is round every time I place an object onto it or walk around it. Thus, while skepticism may succeed at the theoretical level, it is virtually impossible at a practical level. The second problem is that, even if we could suspend all of our beliefs, many of life’s events would still make us unhappy. Like Sisyphus, I can still be bored to tears with my assembly line job even if I doubt that the factory actually exists. Like Job, I can still suffer enormously if my family dies in a tornado, even if I doubt whether my family actually exists. We experience many painful emotions independently of our belief convictions, and skepticism has no solution for those.

**Cynicism and Defying Convention.** Some years ago a music festival was launched called Lollapalooza, which traveled the country attracting large crowds of young people. While the musical groups varied, many were in the crude and abrasive Punk genre—championing a garage band sound, often with instruments out of tune and vocals off pitch. One band included a “percussionist” who grinded away on a chunk of sheet metal with an industrial disk sander. The festival was so successful that it became a yearly event and several non-musical performances were added, including a television-smashing pit. Most bizarre was a circus sideshow in which one performer ate broken glass, another impaled his cheeks with long skewers, and another lifted heavy weights from body piercings. With its growing notoriety, Lollapalooza became a symbol for a burgeoning youth counterculture that was frustrated with sterile social expectations and rebelled against prevailing values.

Many of our conceptions of human happiness are rooted in traditional social expectations, such as how we should dress, what counts as good music, what we should find entertaining, how we should view authority figures. These expectations are not only restrictive, but often wholly misguided. One solution to the question of life’s meaning is to challenge cherished social conventions and through this act of defiance awaken a broader appreciation of life’s possibilities. The social rebelliousness of recent youth cultures is in many ways an embodiment of the ancient Greek philosophical school of Cynicism. The aim of that ancient movement was to show contempt for traditional social structures and values—such as power, wealth and social status. By doing so we would rethink the influence that civilization should have on our lives, open ourselves up to a more direct connection to nature, and thereby become “citizens of the cosmos.” Cynicism was more a way of life than an exact philosophical theory, and its defenders were notorious for their shocking behavior.

A case in point is Diogenes of Sinope (c. 410-320 BCE), who lived as an impoverished beggar in protest of the increasingly lavishness lifestyles of his peers. As a young man, he was exiled from his home town for defacing coins, which were symbols of economic power and political authority. When someone asked him how he felt about his exile, he replied, “That was how I became a philosopher, you miserable fool!” Hunting for a place to stay, he took up residence in a large barrel, which quickly
C. Western Religious Solutions

For millennia religious traditions around the world have assumed the task of explaining the meaning of life. For whatever woes we have, there is some spiritual explanation that aims to redirect us. The dominant religions in Western civilization are Judaism, Christianity and Islam, and among these we find some common perspectives on life’s meaning.

**Having Children.** One of the more famous stories from both the Jewish Bible and Muslim Koran is that of Abraham, a nomadic herdsman who longed to have children in spite of the fact that his wife was infertile. Monitoring Abraham’s situation, God offered him a deal: if he accepts God as his deity, then his descendants will be as numerous as the sand of the sea. Abraham agreed, he had his children as promised, and ultimately became the father of both the Jewish and Arabic people. While there isn’t anything uniquely “religious” about the desire to have children, many faith traditions list this as one of their top religious duties. Conservative Judaism is a case in point. Unlike other faiths that emphasize life after death, Judaism stresses the world right here and now; as one of their sacred texts states, “Better is one day in this life than all eternity in the world to come.” Reproduction is a way of achieving a type of immortality in the present world. I die, but my name, my legacy, and my family history live on through my children.

Some religious philosophers argue that God implants instincts in human nature to help guide our conduct on earth, one of which is the drive to procreate. A more secular understanding of this crucial urge is that it is the result of blind evolutionary forces which keeps animal species like ours from going extinct. Regardless of whether the desire to procreate originates from God or blind evolution, though, it’s a fact of human nature that when we reach a certain age, we have a compelling desire to have children. When we succeed, we magically gain fulfillment and a larger sense of
purpose beyond our individual lives. On the other hand, failing to have children sometimes results in a sense of incompleteness and, in old age, loneliness. To combat this, childless couples frequently transform their pet dog or cat into surrogate children, and lavish love and attention on them to a degree that others find comical. Sometimes it works, other times it doesn’t. So it seems that nature rewards us when we answer its call to produce offspring, and punishes us when we don’t.

While procreation might very well give us a purpose beyond our individual selves, is it a cure-all for the problems of life’s meaning? Perhaps not. First, having children invites a new set of miseries for parents. There is the need to cut back on one’s most cherished private leisure activities to make time for the exhausting task of child-rearing. There are the constant worries about physical dangers to our children, from poorly designed highchairs to automobile accidents. There is the endless battle to block the bad influences of sex, drugs and violence in the media and schools. There are the inevitable clashes with children during the terrible twos, the rebellious teens, and all years in between. We also suffer along with our children when they are harmed or upset—as reflected in a recent expression that parents are only as happy as their saddest child. Marriages often suffer as a direct result of children, sometimes because of a decline in marital intimacy as privacy becomes impossible, other times because of fights over who should do which child-rearing chores. When things end in divorce, the presence of children can lead to vicious and all-consuming custody battles. In effect, having children involves exchanging Sisyphus’s problems for Job’s.

Second, there are limits to how procreation solves Gilgamesh’s problem of human mortality. It’s an exaggeration to say that we gain immortality through our children who will outlive us by perhaps only 25 years. Our grandchildren might extend this by another 25. Generations beyond that, though, will consist of people that we’ll never know, and who will have no memories of us apart from what is conveyed in a few faded photos. The illusory nature of this kind of immortality may become more evident when our children leave the nest, take on lives of their own and become almost strangers to us. We’re once again on our own to find meaning, this time, though, while our health declines and our friends die one after the other.

**Life after Death.** With the limited success of procreation as a cure for life’s anguish, religion offers a backup plan: finding meaning in this life through the prospect of immortality in the next. Most faith traditions present some account of life after death. While the details vary, the core notion is that the essential part of my conscious identity survives the death of my body in a more perfect state of existence. I might exist in a three-dimensional form that resembles my current shape, but is constructed from a more flawless substance. Alternatively I might exist as a purely spiritual thing that takes up no three-dimensional space. In any case, the real me lives on after my body dies.

With a single blow, the idea of life after death attacks all three classic problems of life’s meaning. Most obvious is its solution to Gilgamesh’s problem of mortality. The fact is that we never really do die. Upon the death of my body, my true self is released from its physical shackles and continues in another realm. I may not at first enthusiastically embrace the idea of physical death, which is understandable, like my reluctance to throw away an old comfortable pair of jeans for a new pair. But when I fully grasp that my real self will be preserved through this transformation, my worries about death should fade. Life after death also addresses Sisyphus’s problem of life’s pointlessness. My life’s activities may seem futile to me right now, but that’s because the physical world that I currently live in is imperfect. My efforts on earth are really just a preparation for the world to come, and as long as I keep that in view, life right now has a very clear and important point. Finally,
afterlife is consistently strong enough to counteract the problems of life’s meaning. Sometimes it is, sometimes it isn’t.

Suppose that you can get past this first obstacle and you firmly believe in an afterlife—at least most of the time. The next obstacle is having confidence that you’ll actually get there. While belief in an afterlife may be widespread, paths to getting there are as numerous as the many world’s religions themselves. Again, we must set aside the question of which, if any, is actually the true path. The more pressing question is how confident you are that you’ve picked the right one. Do you have nagging doubts that maybe the denomination across the street is a better gamble than yours? Further, most religions set out pretty tough requirements for entrance into an afterlife—being morally upright, regularly following specific religious rituals, devotion to specific religious founders, earnestly believing a long set of doctrines. Are you sure that you dotted all of your i’s and crossed all of your t’s on that front? While you might still enter the afterlife, a serious oversight on your part might send you to hell rather than heaven. Insecurity about the fine print might burden believers with more unhappiness in this life, rather than relieve the worries that they already have about a meaningful life in the here and now. At best, the hope of life after death will have limited success in giving meaning to life, and, at worst, it may add to our earthly torment.

Furthering God’s Kingdom. During the collapse of the Roman Empire, the Christian philosopher Augustine (354-430) wrote a treatise titled The City of God in which he described two “cities.” On the one hand there is an earthly city, which represents a way of life that is driven by self-love and contempt for God. People of the earthly city are unhappy and experience despair since earthly notions of self-love are so distorted and misguided. On the other, there is a heavenly city which is a way of life that glorifies God and is ultimately achieved.
in the afterlife. While on earth, though, followers of the heavenly city live a God-centered life and work to advance God’s kingdom. This defines who they are, and gives a meaning to their lives which followers of the earthly city cannot experience. One Christian denomination that exemplifies this devotion to God’s kingdom is the Latter Day Saints—better known by their nickname the Mormons. Upon leaving high school, every young Mormon man and woman is expected to serve as a missionary for two years, often taking them to the far corners of the world. During this time they abstain from the leisure activities of watching movies, playing sports, and listening to popular music. Their single focus is to spread the message of God and baptize new believers. Through their devotion they become connected with a higher purpose which gives a special meaning to their lives.

Regardless of the denomination, there are several common features that these religious missions exhibit, which make them larger-than-life experiences for believers. First, these are typically group-efforts among a community of believers, rather than simply isolated campaigns of individual people. It is often this connection with a larger group that gives believers a sense of belonging that they would not otherwise have and, thus, enhances their sense of life’s meaning. Second, there is devotion to a firm and sacred set of beliefs about God’s role in human affairs. While theories about the nature and existence of God are a dime a dozen, not just any view of God will do. Leaders within these religious traditions formulate precise doctrines, and believers pledge exclusive devotion to them, thereby rejecting the views of rival religious groups. By embracing these sacred doctrines, believers see themselves as embarking on a higher mission from God and not just participating in routine human-created social activity. Third, participating in this higher mission involves self-sacrifice. Furthering God’s kingdom is no easy task: it is financially costly, time consuming, and mission efforts are commonly met with brutal opposition. Yet, by enduring these hardships, believers feel a special accomplishment when they make progress.

While the notion of “furthering God’s kingdom” is distinctly religious in nature, there are many non-religious groups that similarly try to transform society through some moral or political ideology. Like religious missions, these involve group efforts among like-minded people who are devoted to a precise higher calling and willingly endure hardship. Some of these social causes are preserving the environment, eliminating poverty, defending political freedoms, ending minority oppression, or creating global harmony. Devotees to these ideologies often gain a sense of life’s meaning similar to that of their religious counterparts.

Whether religious or secular, there is a serious price to pay when devoting oneself to a higher mission, namely, conformity. For a group to speak with a single voice, individual members must give up much of their private identities and follow the direction of the larger collection. There is little room for dissenting opinions about the precise nature of the higher mission: this is firmly fixed in the group’s sacred doctrines and enforced by their leaders. Many believers are content to uncritically follow the directives of their traditions, but others are not. Ex-members of conservative religious groups regularly describe how restricting life was for them and how their leaders used various intimidation tactics to keep them in line. Their leaders, in turn, dismiss the disaffected members as mere trouble makers. The clash between the individual and group becomes more serious when the group’s missionary tactics are morally questionable, such as launching smear campaigns against rival religious groups. Loyal members comply, outspoken critics are shown the door. While furthering God’s kingdom may help give meaning to the lives of some people, it will be less effective for nonconforming “troublemakers.”
D. Eastern Religious Solutions

Religions of Eastern Civilization include Daoism, Buddhism, Hinduism, and perhaps a dozen other traditions of varying sizes. Eastern religions have advocated all three of the above Western solutions to life’s meaning with their own regional twists. In addition to these, though, we find solutions grounded in the more unique philosophical elements of the Eastern traditions themselves. Again, it is important to understand that each religion—Eastern or Western—has its own long and elaborate tradition of beliefs and worship practices that give meaning to believers’ lives. We will look at only three highlights here.

**Daoism and the Way of Nature.** The Daoist religion emerged in China about 2,500 years ago, and its principal message is that of the *Dao*—the Chinese word for “way” or “path.” More precisely, it is the path of nature itself, which creates and guides everything we see. The Daoist solution to life’s meaning is picturesquely presented in a classic tale. One day a prince stopped by to see his cook who was in the process of cutting beef. All of the cook’s motions were like a harmonious ballet as he placed his feet, moved his knees, heaved his shoulders and brought his knife down on the meat. “You have very admirably perfected your art” the prince said. The cook laid down his knife and explained, “I follow the Dao, which is more important than any other skill. Many years ago when I began cutting meat, all I saw was a large chunk of flesh, which I chopped away at. In time I noticed the natural crevices in the meat and, in a spirit-like manner, allowed my knife to glide through them with ease. By doing this I avoided tough ligaments and large bones. An ordinary cook changes his knife every month because he hacks. A good cook changes his every year because he cuts cleanly. I’ve been using this knife now for nineteen years.” The prince responded, “Amazing! By hearing you speak of your craft, I’ve learned how to tend to my life!”

The cook’s message is that we should live in accord with the flow of nature, and not aggressively go against it. Picture a stick floating down a river. When it bumps into a rock, it doesn’t bash its way through the obstruction; instead, it gently moves around it and continues down its course. Daoism has a range of specific recommendations for how we should tend to our lives. For example, we should abandon needless rules of law, morality, and etiquette and instead spontaneously follow the simple inclinations that nature has implanted in us. When we are hungry, nature will direct us to acquire food. If other people are hungry, nature will direct us to assist them. We should even avoid expanding our knowledge through study since this will obstruct the wisdom that nature has already placed within us. By following the Dao, our entire social environments will be transformed. Gone will be the hustle and bustle of big cities, our reliance on intrusive technology, and the endless conflicts between each other. We will instead live more tranquil lives in natural surroundings, and work more directly with nature to meet our immediate needs.

Daoism also has definite suggestions for dealing with life’s woes. If I become gravely ill, I should recognize this as part of the natural cycle of things from growth to decay. The prospect of dying itself should not agitate me since from my raw elements nature will bring forth new life in the continuing cycle of birth and death. By understanding and experiencing the Dao, I see my place in the natural course of things, yield to its power, and peacefully accept whatever happens to me.

It’s hard to be critical of a philosophy whose central theme is to return to nature—an intuition that resonates within many of us. But Daoism relies on an extreme conception of human nature that is difficult to accept. Let’s grant that humans are products of natural forces and we are ultimately at the mercy of natural cycles of growth and decay. Still, humans come into existence with very few natural inclinations to guide us through life. Our survival skills
have been honed through thousands of years of trial and error, and passed down from one generation to another. Without this pool of acquired knowledge to draw from, even the simplest task of finding our next meal would be insurmountable. Nature’s unforgiving message to the human race has been “Good luck in finding your own way: if you fail, you die.” There are indeed important lessons that we can learn from observing nature at work, but the most valuable of these—successful farming, cures for diseases—result from careful and even scientific analysis. Our actual lot in life does not seem to be as passive as Daoism suggests, and any happiness we experience must be achieved while we actively acquire the knowledge that we need to survive. Thus, there is little opportunity to peacefully glide through life as Daoism recommends, and its solution to the problem of life’s meaning looks a little impractical.

**Buddhism and Extinguishing Desire.** The Buddhist religion was founded around 500 BCE in India by a man named Siddhartha Gautama—later given the honorary title of “Buddha,” which means enlightened one. After many years of attempting to achieve enlightenment through traditional religious paths, the Buddha worked out an approach that he felt answered the fundamental question of life’s meaning. He encapsulated his position in the **Four Noble Truths**. The first truth is that life is suffering. From the moment that we’re born until we die, everything we do involves suffering—physical pains, emotional traumas, endless frustrations and disappointments. While Job’s problem of suffering was sparked by specific tragedies, Buddha maintained more generally that suffering is an integral part of our daily lives and is the obstacle we face in our life’s quest for meaning and happiness. The second noble truth is that the source of all suffering is desire. We crave almost anything that might bring us pleasure—sensuality, personal opinions, cherished traditions—and these yearnings become so intense that they rise to the level of addictions. The third is that the cure for suffering is the elimination of desire. If suffering is caused by desire, then it stands to reason that by eradicating desire we thereby end suffering. The fourth is what he calls the **eightfold path**, which is a series of eight techniques for eliminating desire. The specific paths involve the proper cultivation of our understanding, thought, speech, action, livelihood, effort, mindfulness and concentration.

Buddha modeled the Four Noble Truths after the method used by physicians of his day for treating illnesses, namely, identify the disease and its cause, determine whether it is curable, and then prescribed the cure. This is a perfect way of understanding the issue of life’s meaning: target the problem and offer a solution. Virtually all of Buddhist philosophy and theology is an extended commentary on the Four Noble Truths. But it is the third truth—eliminating desire—that concerns us here. This is the celebrated Buddhist concept of *nirvana*, which literally means “to extinguish.” In essence, we extinguish our desires just as we might blow out the flame of a candle. Extinguishing our desires, though, involves much more than losing our various cravings: I must lose my individual identity and even my self-consciousness as a distinct being. As long as I experience life in the usual way, I will be tainting everything through my private, self-indulgent identity. By crushing my identity, I crush all the suffering that I’ve created through my desires.

From one perspective, nirvana appears to be the ultimate solution to the problem of life’s meaning. Once the “I” is removed from the equation, there is nothing left to experience life’s misery. There are questions about this solution, though, which Buddhists themselves are quick to raise. First, to truly extinguish my identity, don’t I have to be dead? As long as I remain alive, I will always be experiencing my self-identity. It seems strange to say that the goal of life is to be completely annihilated through death. Second, while most Buddhists feel that nirvana can be achieved while we are still alive, the concept of nirvana-in-this-life is almost impossible to describe,
and very difficult to achieve. The Dalai Lama, one of Buddhism’s great leaders, describes the frustration that many Buddhists experience regarding nirvana:

I myself feel and also tell other Buddhists that the question of nirvana will come later. There is not much hurry. But if in day
to day life you lead a good life, honestly, with love, with
compassion, with less selfishness, then automatically it will
lead to nirvana. Opposite to this, if we talk about nirvana, talk
about philosophy, but do not bother much about day to day
practice, then you may reach a strange nirvana but will not
reach the correct nirvana because your daily practice is nothing.

[Kindness, Clarity, and Insight, Chapter 1]

In short, nirvana is shrouded in mystery, and the best I can do is follow the recommended paths for achieving it, while closing my eyes to what nirvana actually is. While nirvana might be an effective solution to the problem of life’s meaning, it is difficult for us to examine this possibility when we can’t easily put it into words.

Hinduism and the Four Goals of Life. Originating around 3,500 years ago, the Hindu religion is a diverse collection of beliefs and practices that emerged throughout India’s rich history. The religion has many different gods, devotional practices and philosophies, which believers can freely select from—kind of like a religious a la carte menu. Similarly, when it comes to the question of life’s meaning, Hindu tradition does not restrict itself to one simple answer. Rather, it offers four distinct goals of life, which people should embrace in varying degrees during different periods of their lives.

The first goal of life is pleasure in its assorted emotional and physical forms: food, art, music, dance, and even sex. One of the more infamous Hindu writings, the Kama Sutra, is actually a handbook on sexual activity, vividly describing dozens of techniques. We are naturally inclined to pursue pleasures, and in their proper setting it is fully appropriate for us to fulfill our desires. The second goal is material success. Like pleasure, we’re naturally inclined to acquire wealth and power, which not only keeps us from being impoverished but gives us a sense of accomplishment. The goals of pleasure and material success are most fitting for younger couples who are raising families. As we mature, we embrace the third goal, namely moral harmony, which helps regulate our desires for pleasure and success, but also sparks our social responsibility towards other people. The fourth goal is religious enlightenment where believers become spiritually released from the constraints of human life and attain ultimate happiness. This final goal is best pursued when our family responsibilities are behind us and we can go off in seclusion and practice meditation without distraction.

There is nothing particularly original with any of these four goals individually. We find each advocated by different philosophers from around the world, such as Epicurus’ recommendation to pursue pleasure. The unique insight of Hinduism, though, is that we are complex creatures who change over the years, and there is no single goal that will give us meaning at every stage of our lives. It doesn’t make sense to offer a one-size-fits-all solution to the problem of life’s meaning when people are so diverse. But this is precisely what many philosophers have done. For example, the Stoics recommend that we should resign ourselves to fate, and this single formula, they believe, is the sole solution to the problem of life’s meaning. This may be why the question of life’s meaning still seems to demand an answer after thousands of years of attempted answers: we’ve gotten overly simplistic solutions to a very complex problem.

In keeping with Hindu tradition, perhaps we should approach the meaning of life as we would an a la carte menu: we can pick a few solutions now and, when life’s circumstances change, go back to
the menu and pick a few others. While Hinduism suggests four specific ones—each of which is an excellent menu item—we could add all of the other solutions that we’ve discussed so far, Greek, Western religious and Eastern religious. If I ever become a prisoner of war, then I might want to pick the Stoic option. If life becomes too frenzied for me, I might want to pick the Daoist option of following the way of nature. Religiously inclined people can pick the Christian “furthering God’s kingdom” or the Buddhist quest for nirvana, depending on their religious preference. The point is to seek a solution that best addresses a specific life circumstance or problem. The more creative we are in adding items to the menu, the less likely we’ll be overcome by Gilgamesh’s problem of death, Sisyphus’s problem of futility, Job’s problem of suffering, or any other problem that undermines our sense of purpose or happiness.

E. Philosophy and Life’s Meaning

While the subject of the meaning of life is philosophically interesting in its own right, it is also a good introduction to the various fields of philosophical exploration. We’ve seen that questions of life’s meaning lead to puzzles about how we know things, such as the existence of an afterlife, the nature of nirvana, or whether the table in front of me actually exists. Important questions are also raised about the nature of reality—whether God exists and whether there is a spiritual realm beyond the physical one that we see around us. A range of issues emerge about human nature, such as whether my mind can exist independently of my body and whether my actions are within my control. Finally, there are value questions relating to our conduct towards other people and what kind of social structure will best facilitate happiness.

The discussion of the meaning of life in this chapter also illustrates another feature of philosophy: no theory is sacred, and every proposed idea will invariably be followed by criticisms, and these criticisms followed by responses. For example, the story of Gilgamesh suggests that we are burdened by a fear of death. The life-after-death solution aims to address this fear by giving us hope beyond the grave. One criticism of this solution is that even believers occasionally have doubts about the existence of an afterlife. The debate need not end here, though. The believer might counter this with an argument for the existence of an afterlife; the critic might then challenge the validity of the believer’s argument. Hopefully at some point in the debate one side will seem more compelling than another. But even if the discussion appears to end in a stalemate—as it certainly may with many solutions to the problem of life’s meaning—all is not lost. In some cases truth is not necessarily found in the philosophical theories themselves but rather in the critical give-and-take surrounding those theories.

Finally, we see from the discussions in this chapter that there are different ways of addressing philosophical issues. Some are more scientific, drawing on tangible experience and the contributions of researchers. Others are more introspective, relying on private reflections about the human thought process. Still others rest on pure logic, such as whether two positions can be consistently held at the same time. Philosophy also developed different forms of expression that communicate different content. While many philosophical works are in the form of systematic treatises, philosophy is also expressed in the form of dialogues, meditations, poetic expressions, diaries, confessions, fragments, commentaries, geometric structures. Philosophers have pretty much used every style of written communication to express their ideas. In every case, though, philosophy aims to challenge us to rethink our old assumptions. While philosophy involves a certain amount of mental gymnastics, the rewards are well worth the effort as we begin to see issues in new ways.
For Review

1. What does the story of Gilgamesh suggest about the meaning of life and what is Heidegger’s view of death?
2. What is Camus’ analysis of the Sisyphus story?
3. What does the story of Job suggest about the meaning of life?
4. How would Epicureans respond to criticisms of their solution to the problem of life’s meaning?
5. How would Stoics respond to criticisms of their solution to the problem of life’s meaning?
6. How would Skeptics respond to criticisms of their solution to the problem of life’s meaning?
7. How would Cynics respond to criticisms of their solution to the problem of life’s meaning?
8. What objections could be raised against the “life after death” solution to life’s meaning?
9. What objections could be raised against the “furthering God’s kingdom” solution to life’s meaning?
10. How is the Daoist solution to the problem of life’s meaning comparable to that of the Stoic?
11. What are the Buddhist Four Noble Truths, and what is nirvana?
12. [How do the four Hindu goals of life parallel ancient Greek doctrines?]

For Analysis

1. Defend Heidegger’s view of death against the criticism of it presented in the book.
2. Defend Camus’ solution to the problem of futility against the criticism of it presented in the book.
3. Write a dialogue between an Epicurean and a Stoic regarding pleasure and resigning oneself to fate.
4. Criticize the “life after death” solution from the Taoist perspective.
5. In a quote in the book, the Dalai Lama argues that nirvana is best achieved through day to day moral conduct. Using examples, explain his point.
6. Criticize the Hindu four goals from the Stoic perspective.

REFERENCES AND FURTHER READING

Works Cited in Order of Appearance

Further Reading


CHAPTER 2: GOD

A. The Nature of God
   Personalness, Goodness and Gender
   Power and Separateness
B. Arguments for God’s Existence
   The Cosmological Argument
   The Design Argument from Analogy
   The Design Argument from Probability
   The Ontological Argument
C. Criticisms of Religious Belief
   Belief in Miracles
   Psychological Theories of Religion
D. The Problem of Evil
   The Argument
   Possible Solutions
   The Free Will Defense
E. Faith and Reason
   Pascal: Wager on Belief in God
   James: The Right to Believe
   Plantinga: Rationally-Foundational Belief in God
F. Religious Pluralism
   Four Options
   The Problem of Conflicting Doctrines
   The Problem of God’s Inaccessibility

For Reflection

1. One of the main attributes traditionally ascribed to God is that he is all-powerful. What might be some others?
2. Can God’s existence be proven?
3. How, if at all, can we determine whether a reported miracle is genuine or a fraud?
4. Are some instances of human suffering so bad that they cast doubt on the existence of God?
5. Is it reasonable to believe in God without any evidence whatever?
6. Is there one true religion, or are all world religions connected at some level?

There is a religious organization called “Elvis Underground: The Church” whose mission is “to continue and further the Work of Elvis Presley, The King, toward uniting all species in universal rhythm and harmony.” It holds that Elvis is the most fundamental material in the Universe, and among its Commandments are “Don’t be Cruel,” “Love me Tender,” and “Viva Las Vegas.” That is, we should eradicate cruelty from one self and the universe, love with tenderness, and take risks in life. Although Elvis Underground is more of a tribute to Elvis than a real religion, it nevertheless raises the same kind of philosophical questions that all religions do. What kind of thing is this Elvis principle and how do we know that it exists? Are we irrational if we accept Elvis on the basis of faith without proof? Is Elvis Underground a better religion than others? Philosophy of religion attempts to answer these questions, but without giving preference to any specific faith such as Christianity, Hinduism, or even the Elvis principle. What is important for the philosopher is the abstract concept of God and religious belief, not a particular tradition. While particular traditions may sometimes serve as illustrations, the focus is typically on larger problems.

A. The Nature of God

Philosophers of religion throughout history have been especially interested in clarifying the nature of God—what sort of being he is and what his chief characteristics are. The notion of “God” means different things to different people, and the sky is the limit for our speculations on the subject. The Elvis Underground church lists the attributes of its deity in the following creed:
We feel that Elvis is King. We feel that Elvis is the most fundamental material in this Universe and all others—the foundation upon which all else is built, the fiber that binds all matter and energy, and the catalyst for all action, reaction, and interaction. We feel that Elvis is all good, and all loving, and that by extension, so is this Universe and all others. We feel that limitless goodness and love make Elvis all powerful. We feel that Elvis infuses and embraces this Universe and all others with rhythm and harmony.

Philosophical discussions about God’s nature have typically taken place within the context of conventional religious traditions. Western religions—specifically Judaism, Christianity and Islam—share a core set of views about God, and scholars sometimes refer to this being as the theistic God. The concept of the theistic God has been the focus of much discussion about divine attributes, which we will explore here.

“Personalness,” Goodness and Gender. There isn’t any master list of attributes of the theistic God that everyone agrees on. Some philosophers argue that God has one single attribute, others maintain that he has dozens. Perhaps the most basic attribute is that God is personal: he is conscious and rational, and has the capacity to communicate with other conscious creatures. When and how he chooses to communicate with anyone is another matter, but what is important, according to theists, is that God at least has that ability. By stating that God is personal, the believer is ruling out the possibility that God consists merely of unconscious energy, blind forces of nature, abstract laws of nature, or some other powerful yet insensible component of the natural world. While God may be radically different from humans in other ways, according to theists it is the attribute of “personalness” that makes him most like us. The appeal of a personal God is evident: it presents us with a powerful being who resembles us in important ways and with whom we can communicate in times of distress.

They key problem with the attribute of personalness is that it exposes the concept of God to the charge of anthropomorphism—that is, that the notion of a personal God is too human-like. Suppose that I insisted that God literally had white hair and a beard, just as Michelangelo depicted God in the famous Sistine Chapel painting. You’d say that my concept of God was misguided, perhaps by my desire for an emotionally soothing image of the divine being. But when we survey the attributes that we commonly give God, such as personalness, we discover that many look suspiciously human, which suggests that we just made them up ourselves. The Greek philosopher Xenophanes (c. 570–c.478 BCE) was one of the first to make this point:

If oxen and horses or lions had hands, and could draw with their hands, and execute works of art as people do, horses would draw the forms of the gods like horses, and oxen like oxen, and make the gods’ bodies in the image of their various kinds.

The general problem of anthropomorphism is also expressed in the popular statement that “man created God in his own image”—a reversal of the Biblical notion that God created man in his image. Thus, as comforting as the notion of a personal God is, according to critics, it may be more an expression of human wishes than the reality behind God himself. This is particularly so when we depict God as having human-like emotions. Feelings like love, hate, joy, unhappiness all are the result of physiological activities within our bodies. Since God has no physical body, he can’t have emotions as we do. Similarly, since God doesn’t have physical eyes or ears, he doesn’t perceive the physical world as we do. Theists thus face the difficult task of maintaining the personalness of God without turning him into a giant human being.

Another crucial personality feature of the theistic God is the attribute of perfect goodness, or omnibenevolence. The idea of a
purely evil God has little appeal for anyone—except perhaps for people who are themselves consumed with evil. The concepts of God and perfect goodness seem to go hand in hand, and it is another feature that gives believers comfort, especially in times of distress. The notion of divine goodness includes two distinct concepts: compassion and justness. God is compassionately interested in relieving the world of suffering and enabling people to obtain happiness in this life or in the next. At the same time, though, God has a clear vision of moral justice which he himself follows and commands people to obey. While the notions of compassion and justness may be compatible with each other, there is the possibility of some tension. This is especially so when God considers what to do with people who do not obey his moral commands. While God may want to be compassionate towards everyone, justice may require him to judge and punish immoral people.

A third central issue concerns God’s gender. When many believers think of God they invariably visualize the divine being as either male or female. The common choice among theists is male since this is how God is depicted in the scriptures that have shaped the great religions of Western civilization. But there are a few positions on this somewhat sensitive issue. There is first the staunch pro-male position which holds that scriptural depictions of God’s gender should be taken literally. This, proponents argue, fits with the fact that, throughout human history, men are the ones who have held positions of political, military and economic power. Men have also been the principal architects and builders of machines and structures. So it makes sense to think of God as male insofar as he is the ultimate creator and ruler of all that we see. On the pro-female side, some argue that the creative process of the world involves the growth and nurturing of living things, and fertility is traditionally a female attribute. Feminist theologian Rosemary Radford Ruether writes,

God is not a “being” removed from creation, ruling it from outside in the manner of a patriarchal ruler; God is the source of being that underlies creation and grounds its nature and future potential for continual transformative renewal in biophilic [i.e., life-loving] mutuality. [Women and Redemption]

Many non-Western religions have recognized this and thought of the creator as a nurturing Goddess, rather than as a male God who likes to build intricate things and then blow them up. According to defenders of this position, even if Western scriptures routinely refer to God as he, we can still modify our conception of the divine being’s gender and bring it in line with this idea of a nurturing Goddess. The middle position on this subject is that God is gender-neutral, and it is our gender-laden language that forces us to improperly identify God as a he. By imposing gender on God—whether male or female—we make God overly human, and thus devalue the very concept of the divine being. The idea of the theistic God is flexible enough to be cast in either a male, female or gender-neutral way. In the end, though, questions about God’s gender will largely be determined by one’s religious tradition, regardless of philosophical debates on the subject.

**Power and Separateness.** The attributes of personalness, goodness, and even gender are important for how believers relate to God in their private devotional lives. Other divine attributes, though, are more abstract in nature and are explored by philosophers because of the conceptual puzzles that they generate. Among these are that God is all-powerful (or omnipotent); that God is all-knowing (or omniscient); that God is all-present (or omnipresent); that God is timeless (or omnitemporal); that God is separate from the world. We’ll look at two of these: all-powerfulness and separateness.

Perhaps the most insisted upon attribute of God is that he is all-powerful; the idea of a God that’s anything less than all-powerful does not make much sense. There are two ways that this concept
can be understood. The first is that an omnipotent God can do anything at all, regardless of how counterintuitive the task may seem. Not only can God do colossal physical things, like create the universe, he can even perform seemingly impossible tasks, like making $2+2=5$. While this notion of omnipotence aims to amplify God's greatness, it also creates problems. Can he destroy himself? Can he create a being more powerful than he is? Can he exist and not exist at the same time? If God can do anything, then he should be able to perform each of these, regardless of how bizarre they seem. The basic problem is encapsulated in this popular riddle: Can God create a rock so large that he can’t move it? Suppose that we answer “yes, an all powerful God can indeed create a rock so large that he can’t move it.” We now have conceded that there is a task that God cannot perform, namely, moving the rock that he just created. Suppose instead that we answer “no”; we’ve now just conceded that God can’t do something, namely, create the rock. So, no matter how we answer this question, it looks like there’s something that God cannot do.

The second view of divine omnipotence tries to avoid this problem: an omnipotent God can do anything that is logically possible, but he cannot perform logically contradictory tasks. God can, for example, create the world, since this is logically possible; he couldn’t, however, create the earth and not create the earth at the same time. God also couldn’t perform the logically contradictory task of making $2+2=5$. According to defenders of this position, no possible being could do these things, and so it’s not really a restriction of God’s power. This view of omnipotence solves the problem of the rock: it is logically impossible for an all-powerful being to create a rock so large that he can’t move it. Other problems are resolved in the same way: it is logically impossible for God to do things that are incompatible with his very nature, such as God as an eternal being destroying himself.

So, which of these two notions of omnipotence should theists adopt? If you think that logic is a permanent and unchanging law of the universe, then you’ll go with the second conception of omnipotence. It respects the inviolable nature of logic and has God work within its constraints. On the other hand, if you’re not too impressed with logic and think it’s just like any other changeable thing in the universe, then you’ll go with the first conception of omnipotence. It puts God in a position of authority over literally everything, including logic.

A final abstract attribute of the theistic God to consider is that he is separate from the world—that is, the world that we see around us is technically not part of God himself. It is the creative work of God, but not part of his identity. The critical point is that God is a distinct person—just as I am and you are—and his identity does not mix with other stuff in the universe. The attribute of separateness is an integral part of Western notions of God. But Eastern views of God largely reject the attribute of separateness. Instead, they hold a rival theory known as pantheism, the view that God is identical to nature as a whole. According to pantheism, if we drew a circle around everything in the entire universe—physical and spiritual—the contents of that circle would be God. There are many versions of pantheism, but the critical point here is that God literally dwells within everything: rocks, plants, animals, and, most importantly, each human being. God has no distinct identity and, stated boldly, everything is God. Since traditional Western theists are committed to the attribute of separateness, and Eastern religions to the rival pantheism, there is a major conceptual rift between the two groups.

B. Arguments for God’s Existence

It’s one thing to clarify the notion of the theistic God with a list of attributes. It’s another to demonstrate that such a being exists. Since the middle ages many philosophers have believed that they could
do just that. One website boasts over three hundred proofs for God’s existence, and includes this “argument from guitar mastery”:

(1) Jimi Hendrix is God;
(2) Therefore, God exists.

More traditional proofs for God’s existence follow several strategies, four of which we will examine here: the cosmological argument, the design argument from analogy, and the design argument from probability, and the ontological argument.

The Cosmological Argument. One of the more common proofs for God’s existence is the cosmological argument, which attempts to answer a simple question that even children ask: where did everything come from? The world, it seems, consists of a chain of things that come into and then go out of existence; an oak tree dies, but a new one emerges from an acorn dropped by the first, and the cycle continues. What started it all to begin with? The earliest versions of the cosmological argument maintained that the causal chain of events cannot be extended back through time endlessly, so there must, then, be a first cause to the series, namely, God.

Philosophers in later times rejected this argument since it relies too heavily on the limits of the human imagination. Just because we can’t imagine an infinitely long causal chain of events doesn’t mean that it’s impossible. The real question is whether the concept is inherently contradictory. While it may be difficult for us to picture it, the concept of a universe that has existed for all eternity isn’t logically contradictory, so we can’t rule it out as a possibility. [Besides, if God is the cause of the universe, then what causes God? If God does not need a cause, then why should we think that the universe does?]

A more sophisticated version of the cosmological argument was offered by German philosopher Gottfried Willhelm Leibniz (1646–1716). Leibniz begins by granting that the causal chain of events in the world around us traces back through time to an infinitely remote past. Specifically, the origin of each object within that chain is explained by the object that caused it: the origin of an oak tree is explained by its parent, that oak tree by its parent, and so on infinitely. However, Leibniz argued, God is still needed as an explanation for the fact that the entire infinite series of causes exists at all. In other words, why does something exist rather than nothing? The explanation must then reside outside of the series; that is, it must be a necessary being that isn’t part of the total chain of dependent things. More formally, the argument is this:

(1) The world contains an infinite series of dependent objects.
(2) The explanation of the series is either within the series itself or a necessary being outside that series.
(3) The explanation of the series cannot reside in the series itself, since the very fact of the series’ existence would still need an explanation.
(4) Therefore, the explanation of the series consists of a necessary being outside the series.

Scottish philosopher David Hume (1711-1776) was a staunch critic of various proofs for God’s existence, including the cosmological argument. The central flaw in the above argument, he argues, is that the idea of the “whole series” is only a mental abstraction that doesn’t have any real existence. Once we’ve adequately explained each particular thing in the infinitely long chain, such as the origin of each oak tree from its parent, our job is done and there’s no “whole series” that’s left to explain. Imagine, for example, that I have a bundle of pencils in my hand. I place one pencil on the table, and still have the bundle in my hand. I then place the next pencil on the table and still have the
bundle in my hand. However, when I place the last pencil on the table, I clearly have no bundle left in my hand. The notion of a “bundle” was fully accounted for by each particular pencil. In a similar way, Hume says that the existence of the entire series of dependent beings is fully explained by the existence of each thing in the series.

[Also, why does there have to be a reason for the existence of the universe at all? Maybe existence itself is simply absurd.]

**The Design Argument from Analogy**. Perhaps the most intuitive argument for God’s existence is the one from design: the appearance of design in the natural world compels us to believe that there is a divine designer. There are several versions of this argument, the most famous of which is that from analogy. We start by looking at marvels of human ingenuity, a favorite example of which is a mechanical watch. The parts of a watch are crafted with masterly precision, and all the elements work together to fulfill a purpose. It is quite evident that the watch is the product of intelligent design, and could not have emerged through mere chance. Now, when we look to the natural world, we see a similar purposeful design in things—complex elements working together for specific purposes. While parts of outer space seem a bit chaotic, in other parts of the universe design is undeniable. Look at the regular orbits of the planets around the sun. Look at the precise biological design of the human hand, and how all the bones and muscles work together with great dexterity. These surpass the craftsmanship that we see in any watch. By parallel reasoning, then, we must conclude that key parts of the natural world are also the product of intelligent design. The specific analogy employed here is this:

| watch       | . | designed parts and universe |
| watchmaker  | . | designer                     |

Laid out more precisely, the argument is this:

1. Machines such as watches are the products of intelligent design.
2. Parts of the natural world resemble a machine.
3. Therefore, it is highly probable that parts of the natural world are the product of intelligent design.

The success of this argument depends on two issues. First, relating to premise 1 above, we must show how watches are inextricably linked with watchmakers. Every time we see a watch, we must be justified in concluding that it is the product of a watchmaker. **William Paley** (1743-1805), a bold defender of the design argument, persuasively shows that we would be justified in concluding in almost any conceivable situation that a watch was produced by an intelligent being. For example, we could infer the existence of a watchmaker even if we never saw how the watch was created, or if some parts of the watch didn’t work properly, or if we didn’t know what all of the parts were for.

The second issue, which relates to premise 2, is that the theist must show that the natural world sufficiently resembles the craftsmanship of watches. While there are some design similarities between watches and the biological intricacies of human hands, critics of the design argument have contended that the parallels are exaggerated. This is particularly evident when we consider how the theory of evolution provides an alternative and naturalistic explanation of the origin of apparent design in the natural world. What at first appears to be intentional design in the natural world may instead be better explained through blind natural forces. Astronomers have plausible naturalistic accounts of how our solar system formed, complete with its intricate orbits. Biologists have equally plausible accounts of how the human hand developed
through biological evolution. The existence of a watchmaker is indeed the only available explanation for the origins of watches. But, the more plausible that evolutionary explanations of the natural world become, the less similar that watches and human hands seem to be. Thus, the argument from analogy fails because of an important dissimilarity between watches and things like hands, namely, hands can be explained by natural evolutionary processes, and watches can’t.

A final problem relating to the argument from analogy is that even if we concede that the universe is the product of intelligent design, we are not justified in concluding very much about the nature of the creator himself. He may not be a single, all powerful, or all good being. For all we know, Hume says, the world “is very faulty and imperfect, compared to a superior standard; and was only the first rude essay of some infant deity, who afterwards abandoned it, ashamed of his lame performance.” In other words, at its very best, the design argument might show the existence of some divine being or beings, but not necessarily the existence of the theistic God.

**The Design Argument from Probability.** Even in Darwin’s day, many religiously-minded philosophers accepted the theory of evolution, but still felt that the grand design of things was evidence of a divine designer. While the argument from analogy may not be the most effective way to demonstrate this, other strategies are available. One approach, which is particularly popular now, is the design argument from probability. The central idea is that the physical conditions that make life possible on earth are extraordinarily delicate, and their occurrence is more probable under a theistic hypothesis than an atheistic one. Defenders of this position mention several physical conditions that are fine-tuned for supporting life on earth. If the initial big bang as physicists describe it had differed in strength by only the tiniest amount, life on earth would have been impossible. So too if gravity had been a little stronger or weaker. More formally, the argument is this:

1. The existence of life-sustaining conditions is probable under theism, but very improbable under atheism.
2. When considering two competing hypotheses, we should accept the one that offers the most highly-probable outcome.
3. Therefore, we should accept the theistic hypothesis as an explanation of the world’s life-sustaining conditions.

The critic has a response to this argument. Are life-sustaining conditions “very improbable” under the atheist hypothesis as the believer charges? It’s true that if we randomly picked out a planet in the universe, the odds are very slim that it would exhibit life-sustaining fine-tuning. However, the more planets that exist, the greater the chances are that *some* planet will be suited for organic life. There are perhaps two trillion solar systems in this galaxy, and perhaps 500 billion distinct galaxies in our universe. There’s no telling how many distinct universes there are or have been; for all we know, we may be in a long line of universes that have come into and gone out of existence. In any event, many scientists believe that the odds of naturally-occurring conditions are high enough for there to be organic life on other planets throughout the universe, and not just on ours.

According to the critic, the appeal of the probability argument is more emotional than it is a matter of the cool weighing of probabilities. It’s pretty scary to think that life on earth might not have existed if the fine-tuning of things had been off just a little. But our lives are filled with “what if” scenarios. What if I turned left at the stoplight; I would have been killed in an accident. What if I didn’t look in the help wanted section of the newspaper; I wouldn’t have gotten this job. All of these “what ifs” involve subtle
fine-tuning, but we should not push the issue to the point of concluding that a divine designer had his hand in these matters. Fortunate events naturally happen just as do unfortunate ones. The probability argument presents a big “what if,” but, says the critic, perhaps we should just be happy that things unfolded as they did, and leave it at that.

The Ontological Argument. One of the most remarkable arguments of any kind in the history of philosophy is the ontological argument for God’s existence. All of the above arguments for God begin with an observation about the physical world, particularly causal chains of events and the presence of design. It’s like seeing evidence of a divine footprint and tracing it back to God himself. The ontological argument, though, doesn’t begin with any such observation about the physical world. Rather, it is based purely on the concept of God, which, according to the argument, contains within itself the notion of God’s existence. The originator of the ontological argument was a monk named Anselm (1033–1109), who devised it specifically so that he could have a single, stand-alone proof for God that didn’t hinge on extraneous assumptions about the physical world.

By any standard of difficulty, the ontological argument is a challenge to grasp, but the heart of it is this:

1. God is defined as “The Greatest Possible Being.”
2. The Greatest Possible Being must have every quality that would make it greater than it would be otherwise.
3. Having the quality of real existence is greater than having the quality of imaginary existence.
4. Therefore, the Greatest Possible Being must have the quality of real existence.

The key to this argument is the definition of God in premise 1: God is “The Greatest Possible Being”—or in Anselm’s more unwieldy wording, God is “that than which nothing greater can be conceived.” Think of all the great beings that have graced this planet: elephants have great power, Einstein had great intelligence, Gandhi had great moral character. We can conceive of possible creatures, too, that have even more greatness: Superman’s power surpasses that of an elephant. Iron Man’s intelligence surpasses that of Einstein. Charity Man’s moral character surpasses that of Gandhi. Within the spectrum of beings with great qualities, both real and fictitious, the concept of God is that of the Greatest Possible Being. Whatever possible being has the maximum of greatness, that’s what we mean by God. For clarity, let’s abbreviate “Greatest Possible Being” with GPB.

Power, intelligence, and moral character are all great-making qualities; that is, if you have them, you’ll be greater than if you lacked them. Premise 2 tells us that the concept of the GPB includes every great-making quality to its fullest. For example, the GPB’s power is at the highest possible level, more so than Superman’s. Why so? Because if the GPB lacked ultimate power then, by definition, it wouldn’t be the greatest possible being: it could have been greater by possessing ultimate power. So to with qualities like ultimate intelligence and ultimate moral character. Again, if the GPB lacked either of these qualities, it wouldn’t be the greatest possible being. The GPB is sort of like the ultimate superhero insofar as it possesses every great-making quality to its fullest. And, for all we know at this stage of the argument, the GPB is just as fictitious as any superhero. The point is that the very concept of the GPB includes the maximum of every great-making quality.

Premise 3 states that the quality of existence is yet another great-making quality: if you have it, you’ll be greater than if you lacked it. A real superhero, for example, is greater than it would be if it was just a fictitious one. This pulls the GPB out of the realm of
mere superhero fiction and into real existence, and thus leads to the arguments conclusion. That is, the GPB must exist, since if the GPB lacked the great-making quality of existence, then by definition it wouldn’t be the greatest possible being. Reduced to a single sentence, the ontological argument is this: the Greatest Possible Being must exist, since it is contradictory to suppose that it doesn’t—we would in essence be saying that the Greatest Possible Being could have been greater.

Philosophers have made careers out of analyzing and attacking Anselm’s argument. We’ll consider the most famous criticism, offered by a monk named Gaunilo, who was a contemporary of Anselm. Gaunilo argues that if we replace the phrase “Greatest Possible Being” with “Greatest Possible Island,” then we’ll get the absurd conclusion that the Greatest Possible Island exists. Consider, Gaunilo says, the mythical “Lost Island” where everything is perfect. Paralleling Anselm’s argument, here’s how we can prove that the Lost Island exists:

(1) The Lost Island is defined as “The Greatest Possible Island.”
(2) The Greatest Possible Island must have every quality that would make it greater than it would be otherwise.
(3) Having the quality of real existence is greater than having the quality of imaginary existence.
(4) Therefore, the Greatest Possible Island must have the quality of real existence.

Gaunilo concludes “You can no longer doubt that this island which is more excellent than all lands exists somewhere.” His point is that Anselm’s type of argument would show the existence of the greatest possible anything, which means that there is something fundamentally flawed with the logic of Anselm’s argument.

[To this argument Anselm replied that islands and other things are not like God, because they are not perfect by definition. The non-existence of the greatest possible island is not a contradiction in terms; but the non-existence of the greatest possible being would be a contradiction. In response to Anselm, critics point out that even if God’s existence is conceptually necessary, that does not prove that he actually exists.]

In recent times, philosophers have taken a generally negative view of cosmological, design and ontological arguments. Part of the reason is because of critiques by Hume and others. Another reason has to do with the nature of religious belief itself. Believers themselves often recognize that we can’t approach religion as though it was scientific theory that can be proven by tangible evidence, as the cosmological and design arguments attempt to do. We also can’t approach it with a mathematical-like proof along the lines of the ontological argument. Belief in God is a more unique kind of conviction that has few parallels in our human efforts to acquire knowledge, and logical proofs like the above seem misdirected.

C. Criticisms of Religious Belief

While believers have defended the idea of God, sometimes offering proofs, nonbelievers have tried to expose problems with the concepts of both God and religion. Occasionally this takes the form of full-fledged arguments for atheism. More often, though, religious critics pick away at weak links in the theist’s chain of beliefs hoping that it will eventually fall apart. Two prominent areas of vulnerability are belief in miracles and the psychological factors influencing religious belief.

Belief in Miracles. A Christian minister recently claimed to miraculously make infertile women pregnant; some were past menopause, and others had no sexual relations for years. He invited
the women to Kenya where he exorcised demons from them; they then supposedly gave birth and return to their home countries with babies. Reports of miracles are as old as human civilization, and even in modern times are as abundant as ever. Many are evidently frauds. Kenyan investigators, for example, suspected that the minister was part of a child smuggling ring: one Kenyan official stated “We believe in God, but we do not think God works that way.” Are any alleged miracles, though, worthy of belief? David Hume thought not. Hume’s position involves three critical assumptions. First, he defines “miracle” very narrowly as a violation of a law of nature. There are rival definitions of “miracle,” such as unusual events that occur at just the right moment—the wind blowing a vine towards me just as I’m about to slip off the edge of a cliff. Hume does not discuss these. Second, Hume focuses specifically on the credibility of reports and testimonies of miracles. That is, Hume would have me challenge the credibility of the miraculous pregnancies in Kenya that I read about in the newspaper, [and he points out that it is more likely we are mistaken or deceived when we personally witness “miracles.”] He thus questions whether it is reasonable for me to believe reports of miracles, not whether the miraculous event actually took place.

Stated most succinctly, his argument is this: it is never reasonable to believe in reports of miracles since those reports will always be outweighed by stronger evidence for consistent laws of nature. Imagine a balancing scale with two pans. In one pan we place the strongest evidence in support of a miracle. Suppose, for example, that a team of respected physicians examined the women who became pregnant in Kenya, and determined that their pregnancies indeed violated biological laws of nature. The physicians’ testimonies would go into the first pan. In the other pan we place the strongest evidence against the alleged miracle. According to Hume, we need to consider only one factor here: the consistent experience we have of laws of nature. Doors don’t fly off their hinges when we open them. Rocks on the ground don’t mysteriously levitate. When I step on the sidewalk, it doesn’t crumble beneath my weight. The evidence in support of consistent laws of nature is in fact so overwhelming, Hume argues, that it will always overpower any evidence in the other pan, regardless of how credible a testimony in favor of a miracle may be. Hume writes,

It is experience only, which gives authority to human testimony [of miracles]; and it is the same experience, which assures us of the laws of nature. When, therefore, these two kinds of experience are contrary, we have nothing to do but subtract the one from the other, and embrace an opinion, either on one side or the other, with that assurance which arises from the remainder.” [Hume, An Enquiry Concerning Human Understanding, 10]

The wise person, he maintains, will proportion his belief to the weight of the evidence, and the evidence will always be on the side of consistent and inviolable laws of nature, not on the side of miracle testimonies.

Does Hume’s argument succeed? One criticism focuses on Hume’s particularly narrow notion of what counts as a “reasonable belief.” Suppose that miracles never occurred in the past and thus the concept of a miracle was indeed completely contrary to experience, just as Hume claims. Next, suppose that God exists and, for the first time, decides to perform a dramatic miracle in the presence of credible witnesses who then report the event. On the supposition that a genuine miracle did occur, is there any kind of report that would make belief in this miracle “reasonable”? The answer to this question rests on the kind of world view that a person holds, namely, a supernatural world view or a natural one. The supernatural one maintains not only that God exists, but that he directs and possibly interferes with the course of nature. He communicates with people, impacts their lives, and helps them in
times of need. The natural world view, on the other hand, resists all of these claims, placing confidence instead in consistent laws of nature and scientific methods of discovering facts. If I hold a natural world view, then Hume is right: it would never be reasonable for me to believe a reported miracle, [for I would simply revise my understanding of the laws of nature to take into account this unexpected event]. If, though, I hold a supernatural world view, then everything is different. As I try to make sense of life’s strange experiences, I have the option of attributing some of these to supernatural causes. In this way, it might be reasonable to believe a credibly-reported miracle in the context of a supernatural world view.

Hume personally felt that the supernatural world view is rooted in ignorance and superstition, and so he maintained that any belief resting on that world view would be unreasonable. This, though, raises a separate issue about the psychology behind the supernatural world view, which we turn to next.

Psychological Theories of Religion. Religious believers think that the supernatural world view is grounded in divine reality; that is, there is a real God who resides in a real spiritual realm. Non-believers, though, have a different explanation: belief in the supernatural is grounded in human psychology, not in external reality. Two thousand years ago the Roman philosopher Lucretius (94-55 BCE) argued that religious belief is the product of human fear: we encounter frightening natural calamities, can’t find natural explanations of these things, and then “seek refuge by handing everything over to the gods.” Many other philosophers since have adopted versions of Lucretius’ view, and we’ll briefly look at a few to see what they have in common.

Beginning in the 19th century, though, religious critics offered more elaborate and creative psychological accounts of religious belief.

One of these was Karl Marx (1818-1883). For Marx, economic class conflict shapes many of our beliefs, including religion. When Marx wrote, grotesquely rich industrialists dominated society, and made their wealth by exploiting lower-class laborers. If workers wanted food on their tables, they had no choice but to perform tedious tasks for low wages in unbearable conditions. Stripped of their human dignity, workers were offered one avenue of comfort: religion. In church they learned that, while God put them through trials in this life, they should place their hopes in the rewards they’d receive in heaven. In Marx’s words, religion is the opium of the people—a drug which lulls workers into accepting their situations. Marx writes,

Religious suffering is the expression of real suffering and at the same time the protest against real suffering. Religion is the sigh of the oppressed creature, the heart of a heartless world, as it is the spirit of spiritless conditions. It is the opium of the people. [“Toward a Critique of Hegel’s Philosophy of Right”]

Once workers are liberated from the oppressive industrialists, Marx predicts, there will be no further need for religion and religious belief will disappear on its own.

Another influential psychological account of religious belief was offered by Friedrich Nietzsche (1844–1900). According to Nietzsche, for millennia the only world-view available to civilization was the supernatural one—or the “ascetic ideal” as he calls it. This holds that, not only does God exist, but earthly life is worthless, and the only things of value are in the spiritual realm. In recent times, Nietzsche continues, a natural world view has emerged in which philosophers and scientists provide an alternative naturalistic account of things around us. The influence of this rival has expanded to the point that, metaphorically speaking, God is now
dead. That is, the natural world-view has eclipsed the supernatural one and everything it stands for. Nietzsche poetically writes,

Do we not hear the noise of the gravediggers who are burying God? Do we not smell the divine putrefaction? For even Gods putrefy. God is dead! God remains dead! And we have killed him. How shall we console ourselves, the most murderous of all murderers?

While the death of God is in some ways liberating, Nietzsche warns that it is also frightening because it forces us to find an alternative conception of morality.

Perhaps the most famous modern psychological account of religion is that offered by Sigmund Freud (1856-1939). When we were children, Freud argues, we were comforted with the protection that our human fathers gave us, and, from our perspectives, our fathers seemed all powerful. When older we of course learn the truth about our fathers’ limitations. But whenever we experience suffering as adults, we return to our childhood stage of human development and project the concept of a powerful father figure onto the heavens and look to an all powerful God for protection. This again gives us comfort, but it is at the expense of our adulthood. Freud recommends that we abandon the illusory belief in God and reclaim our maturity.

While the positions of Marx, Nietzsche and Freud differ in their particulars, there is a common thread: religious belief is caused entirely by psychological and social forces, and a non-religious view of the world is ultimately preferable. Technically, this is not a disproof of God’s existence or the supernatural realm. But, by tainting the cause of religious belief, it implicitly questions the truth of the beliefs themselves. From the nonbeliever’s perspective, it may be enough of a victory for atheism if believers are kept wondering whether they’ve been psychologically duped.

How might the believer respond? From the religious perspective, there may be nothing wrong with God using psychological factors to direct belief—such as fear of the unknown, the desire for a happy afterlife and the need for a father-figure. After all, humans are psychological creatures. The question is whether these psychological factors are inherently misleading. Marx, Nietzsche and Freud thought they are, and perhaps some of the time we really are tricked into erroneous religious beliefs through these mechanisms. But, according to the believer, it is less clear whether all religious belief is caused by misleading psychological forces. Maybe God really is a father figure, and our psychological projections are right on target. We can’t rule that out as a possibility, the believer argues.

D. The Problem of Evil

A hurricane recently landed on the Gulf coast and demolished everything in its path. One victim, a young woman, stated “Everything I’ve ever worked for is now destroyed. Short of dying, I can’t think of anything worse happening to me right now. I’ve been a good Christian and I pray all the time; I don’t see how God could have let this happen to me.” At some point in our lives most of us raise similar religious questions in the face of our own tragedies. A family member contracts a terminal illness and dies; a friend is killed in a car accident; we lose our jobs; our spouses leave us. In the midst of our despair we may wonder why God would allow these things to happen. He presumably loves us and is in control of things, so what went wrong? Sometimes the despair reaches a point where believers abandon their faith.

Philosophers call this the problem of evil. The notion of “evil” used here simply means suffering, and there are two varieties of
suffering that are central to the problem. The first, often called *natural evil*, is the suffering that takes place through the blind forces of nature. The story of the hurricane is an example of this. The second, often called *moral evil*, is the suffering that takes place because of the willful acts of human beings. A robber breaks into my house and shoots me; a drunk driver crashes his car into mine.

**The Argument.** While the problem of evil begins as a kind of emotional crisis in the lives of believers, religious critics have crafted this problem into an argument for atheism. The basic argument is this:

1. The theistic God, who is all-good, all-knowing, and all-powerful, would prevent evil.
2. Evil exists.
3. Therefore, the theistic God does not exist.

Premise 1 lists the central attributes of the theistic God, and stipulates that such a being would prevent evil, or suffering. Why so? If God is all good, then he would certainly be motivated to eliminate suffering. When discussing the attribute of omnibenevolence, or perfect goodness, at the outset of this chapter, we saw that compassion was part of this notion. That is, God is sympathetic to our suffering and wants the best for us. If God is all-knowing, then he knows where all the suffering is in the world and knows how to root it out. Finally, if God is all-powerful, he has the ability to eliminate the suffering that we experience. Thus, the very nature of God seems to commit him to the elimination of evil.

Premise 2 states the obvious: suffering, both natural and moral, abounds in the world. Since the existence of evil is incontestable, then we must conclude that the problem rests with the existence of the theistic God.

Some religious critics argue that the conflict between the concept of God and the existence of evil is nothing short of a logical inconsistency, and logic demands rejecting the notion of the theistic God. It may be better to see it instead as a tension between the two premises, which makes the existence of God unlikely. There are a couple clear paths to resolving this tension. First, rather than rejecting the existence of the theistic God entirely, we might tweak some of his attributes. For example, we might question whether God is really “all-good.” Perhaps he’s good most of the time, but occasionally lashes out on the world when things become too intense. We might also question whether God is really all-knowing or all-powerful. Maybe God isn’t aware of all the evil that goes on here, or maybe he’s not quite powerful enough to do anything about it. Theists, though, are usually not willing to make these concessions. Built into the notion of God is that he is all-good, all-knowing and all-powerful. Anything less than that, he wouldn’t be God. A second path to resolving the problem would be to deny that evil really exists. Perhaps we’re trapped in a world of illusion in which we erroneously think that we experience suffering. However, whether real or imagined, suffering is suffering. We pity the delusional man who thinks he’s being eaten alive by snakes almost as much as we do the man who really is being eaten alive by snakes. So, the argument from evil looks pretty strong as it stands.

**Possible Solutions.** Many possible solutions to the problem of evil have been offered over the centuries. The term *theodicy* refers to such attempts, which literally means “justification of God,” and suggests that an effort is needed to reconcile the notion of God with the presence of evil. Some solutions to the problem of evil are more plausible than others, and we will look at four of the most influential ones.

The first is that the presence of evil is justified because good often comes out of evil. According to this solution, some kinds of moral goods require the occurrence of evil. For example, major hurricanes
cause billions of dollars of property damage and the loss of many lives. Each time this happens, like clockwork, an army of volunteers travels to the region to help shelter the victims and rebuild their homes. By doing this the volunteers display charity, compassion, humanity, mercy, generosity, kindness, and a host of other good qualities. All of these behaviors are linked to the presence of some human tragedy: you can’t meaningfully be charitable if no one needs your charity. The point, then, is that the world is better off with some evil-dependent goods. Thus, while God is predisposed against all suffering, he allows some of it to make the world better through evil-dependent goods.

The key problem with this argument is that we could have an ample supply of evil-dependent goods with far less suffering. We don’t need enormous disasters like hurricanes to spark charity in people; I would be sufficiently motivated to be charitable to my neighbor if the occasional tree limb fell on his garage and he needed help removing it. When we add up the suffering caused by all crimes, wars, natural disasters, accidents and diseases, it’s hard to see how these are counterbalanced by the charity and other evil-dependent goods that they generate.

A second solution is to re-examine the idea of divine goodness. At the outset of this chapter we noted that God’s attribute of perfect goodness has two components: compassion and justness. It is because of God’s compassion that we presume he would want to rid the world of suffering. Perhaps, though, the believer is over-emphasizing compassion at the expense of justness. That is, maybe God is good in the sense that he demands moral obedience from us, but has limited compassion for our personal suffering. The Greek philosopher Epicurus felt that humans achieve happiness through reducing suffering and enhancing pleasure; perhaps God isn’t an Epicurean. The solution, then, is this: human suffering is compatible with divine goodness when goodness is understood mainly as moral uprightness. While this technically may solve the problem of evil, it would be a tough sell for the believer who takes comfort in the idea of a compassionate God.

A third solution was presented by contemporary philosopher John Hick (b. 1922). He argues that human creation is a developmental process during which we evolve to eventually become a more perfect likeness of God. As individuals, we must strive against hardship to learn, grow, and develop moral character. Everyone knows someone who has never had to face anything difficult and had everything handed to them. But it is through challenges that we mature individually, such as when we overcome social or economic obstacles to get through college and land a good job.

These are also the challenges that enable us to evolutionarily mature as a species over time, and this is the uniqueness of Hick’s theory. Like other life forms on earth, humans emerged first as simple primates and overtime evolved into the sophisticated biological and social creatures that we are today. Throughout this time our predecessors had experienced much suffering, and we will continue to suffer until our evolutionary process is complete. Suffering, then, is simply part of the evolutionary process that drives the moral development of our species.

As innovative as Hick’s solution is, it faces an obstacle: the people who actually suffer the most do not themselves reap the reward of becoming perfect. Instead, their sufferings and struggles are only small steps on a long path that leads to the perfection of human nature thousands of years down the road. But, how much suffering is it reasonable for us to endure right now for the sake of people in the distant future that we cannot even currently envision? And, once humanity has achieved that perfection, how fair is it for those future people to get a free ride at our expense now? If we adopt Hick’s solution, we must recognize that divine justice has little connection
with human notions of justice. This doesn’t rule out Hick’s solution as a possibility, but it does require revamping the notion of divine justice.

The Free Will Defense. The most popular solution to the problem of evil is the free will defense. According to this view, evil is our fault, not God’s, since God couldn’t create a world containing free creatures who would be guaranteed to always be good. Imagine that we looked at the blueprints of all the possible worlds that God could have created. World 1, for example, consists of a single brick floating in empty space and nothing more. World 2 has stars and planets like the present world, except the earth is inhabited by only worms. World 3 is the same, but the earth is inhabited by robotic humanoid creatures with no free wills. They are conscious, but can behave only as programmed. World 4 is the same, but is instead inhabited by truly free human creatures. Which of these is the best? “World 4” we might say. But the problem now is that God has no control over the choices of free creatures. If I have a free will, then it is entirely up to me to decide how I want to behave; even God cannot determine the choices of a truly free creature. If I freely decide to be evil, that’s my choice and, short of striking me dead, God can do nothing about it regardless of how much he’d like to.

God, then, is faced with a dilemma: he could create World 3 and program the robotic humanoids to always be good; or, he could create World 4 and gamble on whether free humans will choose to be good. According to the free will defense, God gambled on World 4: a world of free creatures is better than a world of robotic humanoids, even at the risk of evil. In a sense, the free will defense solves the problem of evil by tweaking the notion of “all-powerful” in premise 1. That is, by creating free creatures, God places a voluntary constraint on his power, thereby handing some power to us to make choices as we see fit. Humans, then, are co-creators of the world alongside God.

In many ways the free will defense seems to be the ideal solution to the problem of evil. It retains the key attributes of the theistic God, acknowledges the existence of evil, and puts the blame on us.

There are three limitations to it, though. First, the free will defense is only an option for those who believe in free will. In centuries past, belief in free will was popular among philosophers; but the advances of genetics and the social sciences has given strength to the view that human actions are all determined by causal forces. If I reject the theory of free will in favor of determinism, then the free will defense goes out the window with it. Second, some theists are not happy with any compromise of God’s power, even if God does this voluntarily. If God could hand over some of his creative power to us, perhaps he could hand over the whole lot to a committee of angels and then go on a long vacation. God’s total control, the theist might argue, is not negotiable. Third, the free will defense is primarily an attempt to explain the existence of moral evil, that is, the evil caused by human choice. It’s less clear how this is an adequate explanation of natural evil, such as floods and hurricanes, which have nothing to do with the free choices of bad people.

[It should be noted, however, that the free will defense ignores the option that God could have created free creatures who are more knowledgeable and virtuous and would thus always freely choose to do good (just as he does). They would not be robots, but more like the blessed in heaven. In this scenario, human beings would choose to be God-like without having to suffer. Furthermore, by saying that God “gambles” on creating a world of free human beings, defenders of the free will position imply that God does not know the results of his creative activity until after he creates—which contradicts his omniscience. Certainly God knows exactly what every person he creates will choose to do, yet he goes ahead and creates them to be the beings who will sin. If the problem of evil is solved by appealing to free will, it thus opens up another problem—namely,
that of how we can be created by an all-knowing God and still be free.]

In the end, then, there is enough substance to the free will defense to explain the existence of some evil to the satisfaction of some theists. And, if the free will defense is taken in conjunction with some of the other defenses above, the problem of evil may be less of a stumbling block than it first appears. [However, even if the problem of evil can be reconciled with the possible existence of God, such reconciliation does not prove that God exists. Indeed, as the Russian novelist Fyodor Dostoyevsky (1821-1881) points out, we cannot rationally prove that God exists, nor should we hope to understand how the existence of God and the existence of evil are reconcilable, because that would make evil something that makes sense and that we are thus willing to accept. Dostoyevsky argues, however, that the existence of evil is absolutely contrary to the possibility of a loving God: that is why belief in God has to be a matter of faith and not reason.]

E. Faith and Reason

So far, the arguments both for and against the existence of God look suspicious, and it appears that strictly rational argumentation cannot settle the dispute one way or the other. Reason, it seems, is neutral on the subject of whether God exists. What are my options now? On the one hand, I could adopt the position of agnosticism, which is neither belief nor disbelief in the existence of God. On the other hand, I could simply believe in God’s existence in the absence of any proof: I could believe through faith. Many religious believers throughout history have gone this second route, arguing that reason has a poisoning effect on a person’s psychological reception towards religious belief. Faith alone, they argue, should guide our paths in religious matters, and not reason. Tertullian (c. 160-220), an early Christian advocate of faith-based belief, made the famous statement that “I believe because it is absurd.” His point is that faith is so distinct from reason that faith is essentially irrational. Tertullian’s position is rather extreme. More moderate defenders of faith-based belief have maintained that while faith is non-rational, it is certainly not ir-rational, and there is in fact some reasonable defense for holding a faith conviction.

Pascal: Wager on Belief in God. The first of these defenders is French philosopher Blaise Pascal (1623-1662). He was familiar with the arguments both for and against the existence of God, and felt that reason was neutral on the matter. But while reason is neutral, Pascal says, we as human beings cannot remain neutral; the question is so important that we must decide one way or another. It’s a kind of wager, he argues, and what is at stake for me specifically is my own happiness—both in this life and the next. Pascal proposes that I weigh the possible consequences on my happiness:

Let us weigh the gain and the loss in taking heads that God exists. Let us weigh these two cases. If you gain, you gain all. If you lose, you lose nothing. Wager without hesitation, then, that he is. [Pascal, Thoughts]

Charted out, here are the stakes of my wager:

<table>
<thead>
<tr>
<th>What if:</th>
<th>God exists</th>
<th>God does not exist</th>
</tr>
</thead>
<tbody>
<tr>
<td>you believe</td>
<td>gain: infinite happiness</td>
<td>loss: very little</td>
</tr>
<tr>
<td>you don’t believe</td>
<td>loss: eternal damnation</td>
<td>gain: very little</td>
</tr>
</tbody>
</table>
According to Pascal, by wagering in favor of God’s existence, I stand a chance of gaining infinite happiness in the afterlife, and losing little or nothing in any event. By wagering against God’s existence, though, I will gain or lose little or nothing. For Pascal, the gamble seems obvious: I should wager that God exists. This is by no means a proof of God’s existence, but rather an indication of the most reasonable choice for me in view of my desire to be happy.

Pascal understood that it’s one thing for me to see the reasonableness of wagering on God’s existence, and another entirely for me to actually believe that God exists. The wager, he explains, is just the start of the belief process. The next step is for me to put myself in a psychological position that is receptive to belief in God through faith. I need to assume the right moral outlook by reducing my private desires and adopting virtues like humility, charity, gratitude, and truthfulness. I then need to follow the example of other believers by going to church and participating in religious traditions. These things won’t make me believe, but, according to Pascal, they will make me receptive to a faith experience which will trigger genuine belief.

What should I think about Pascal’s wager? That is, is it really more reasonable for me to gamble on belief in God? American philosopher William James (1842–1910) felt that the reasoning process of the wager is fundamentally flawed. Rival religions could come along and propose a similar wager for their favorite deity. A Muslim prophet, according to James, could say “I am the Expected One whom God has created in his effulgence. You shall be infinitely happy if you confess me; otherwise you shall be cut off from the light of the sun. Weigh, then, your infinite gain if I am genuine against your finite sacrifice if I am not!” The Elvis Underground church could make a parallel argument for wagering on the divinity of Elvis. As long as reason is neutral on the question, we are not in a position to prefer one religion’s wager over another.

James: The Right to Believe in God. While James rejected Pascal’s wager, he nevertheless defended our right to believe in God on the basis of faith alone when reason is neutral. Key to James’s view is a distinction between different kinds of options that the believer faces. First, according to James the religious options that the believer considers must be live as opposed to dead. A live option is one that the believer might realistically adopt; a dead option carries no psychological appeal. For example, it might be a live option for me to believe in the theistic God, but a dead option for me to believe in the ancient Egyptian God Ra or the divinity of Elvis. Second, the believer’s religious options must be forced, as opposed to avoidable. An option is forced when we face a choice with no middle ground, such as whether I should vote or not vote. Avoidable options have some middle ground; for example if you say “Either love me or hate me,” a third alternative is open to me, namely, to remain indifferent towards you. Religious beliefs, James argues, are at least sometimes forced. If I am faced with the option of either believing in the theistic God or not believing, there is no middle ground since by abstaining from belief I would lose a possible spiritual benefit, just as though I chose not to believe at all. Third, the believer’s religious options must be momentous—that is, of critical importance—as opposed to trivial. For example, usually going on a date to the movies is trivial, but a decision about getting married is momentous. While some people do not take religious belief seriously, many do and in fact find it of critical importance.

James defines a genuine option as one which is live, forced and momentous. He expresses the central point of this distinction here:

Our passiononal nature not only lawfully may, but must, decide an option between propositions, whenever it is a genuine option
that cannot by its nature be decided on intellectual grounds. [“The Will to Believe”]

In at least some cases, religious belief qualifies as a genuine option. That is, when reason is neutral concerning belief in God, and belief in God is a genuine option, then we may rightfully believe on the basis of our emotions. In these situations we can set reason aside and believe on the basis of our hopes, fears, feelings of devotion, and fondness for our traditions. Here is an example of James’s reasoning process. Imagine that John is considering asking Mary out on a date. He looks for subtle signs from her about how receptive she will be, but he comes up with nothing: from his perspective, reason is completely neutral about her romantic interest towards him. For John, asking Mary out is a genuine option: it is something that he’d really like to do (live option); he either asks her out or he doesn’t (forced option); and the potential rewards could be tremendous (momentous option). In this situation, John is fully justified in relying on faith and asking her out. If he doesn’t, then he forfeits all the promise of an important love relationship. According to James, the situation is the same with religious belief, where the potential rewards are religious truth and enlightenment, which could not be gained without acting on faith.

How might the religious critic respond to James’s position? A scientifically-minded person could argue that, when reason is truly neutral on a subject, our duty is to abstain from belief. Suppose that I am investigating whether a specific drug will reduce the risk of heart attack. If the evidence is inconclusive, I shouldn’t go ahead and believe that the drug works, even if holding that belief makes me feel better. It should be no different with questions concerning God’s existence: if the evidence is inconclusive, I should be a religious agnostic. James considers this line of attack and has a response. There are two goals that all inquirers have: (1) pursue truth, and (2) avoid error. The scientist, according to James, places greater weight on the second goal of avoiding error: whatever truths we ultimately arrive at, says the scientist, they should not be at the expense of accepting erroneous views. But James thinks that the situation is entirely different with matters of religion, love, and other non-scientific interests; there our principal goal should be to determine and even establish truth, even if that means choosing error. James’s response is all the more appropriate when we feel that questions about God’s existence can’t be settled with rational proof. It thus seems unfair for the critic to hold religious truths to the same standard as scientific ones.

[That same point is made by Søren Kierkegaard (1813-1855). He notes that the religious person acts “by virtue of the absurd,” without justification. In Kierkegaard’s view, the religious life is based on a “leap of faith” that is not guaranteed or comforted by any rationally guaranteed or universally understandable knowledge of what God wants of us. Instead, it is based on the recognition that we care about our lives because our existence is personal, not something that can be described in terms of objective or provable facts. The commitment to religious belief must thus be understood in terms of subjective, non-universalizable truths on which we are willing to stake our lives.]

Plantinga: Rationally-Foundational Belief in God. James and Kierkegaard give a compelling defense of our right to believe in God through faith when reason is neutral. But while this grants me the freedom to believe in God, it does not allow me to say “I know that God exists” or that “belief in God is rationally justifiable.” This limitation is a problem for many theists who would like to make those stronger claims. Proofs for God such as the design argument have been the usual mechanisms for rationally justifying our knowledge of God’s existence. If those fail, do we then need to abandon all claims to “know” that God exists? Contemporary American philosopher Alvin Plantinga (b. 1932) believes that there is another way to establish knowledge of God’s existence: the
concept of God’s existence is a rational instinct, and not a product of rational proof.

Plantinga’s position rests on a particular theory of how we acquire knowledge called foundationalism. Geometry is an excellent example of this approach. We begin with foundational notions of points or lines, and from these we deduce complex theorems about the shapes of triangles and rectangles. We take this foundationalist approach in other areas of knowledge as well. Take for example my belief that George Washington was the first U.S. President. This is based on more foundational beliefs I have about the trustworthiness of history books and the original eyewitnesses. According to Plantinga, some of our most foundational beliefs are rational—beliefs that flow instinctively from human reason (foundationalists like Plantinga describe such rationally-foundational beliefs as “properly basic.”) Other rationally-foundational beliefs are “the objects that we perceive really do exist,” and “the people I see really do have conscious minds.” Plantinga contends that “God exists” is yet one more rationally-foundational belief. Following this approach, the believer can say “I know that God exists” based purely on a rational intuition, without any rational argumentation.

The central problem with Plantinga’s approach is whether there really are any rationally-foundational beliefs at all. While human behavior is undoubtedly driven by instincts, it is an extraordinarily difficult task to list any instinct-like beliefs that we have, and, on top of that, show that these beliefs are products of human reason rather than human emotion. Further, knowledge seems to rest on an interconnected web of beliefs, and not on a rock solid set of foundational beliefs. Again, my belief that George Washington is the first president is based on views about the trustworthiness of history books. But this in turn is connected with my beliefs about academia, historical method, free speech, and testimony in general. These in turn are based on a thousand other beliefs and experiences, which are interconnected and continually evolving. Geometry, it seems, is a poor model for how we gain knowledge. The whole concept of rationally-foundational beliefs looks suspect.

Plantinga’s theory is driven by a desire to say “I know that God exists” even if proofs for God’s existence fail. But why not just be content with saying “I believe that God exists” without any claim to knowledge? The psychological motivations here are complex. Perhaps part of what’s at stake may be the believer’s level of assurance that there really is an afterlife: our assurance is higher if we know this as a fact and not merely believe it. Part also might be the need to proselytize with conviction: it’s easier to persuade others to join one’s religious tradition if it is backed by knowledge of God’s existence rather than mere belief in God. But belief even without knowledge is a very powerful tool and, as James suggests, is sufficient for directing important decisions in our lives, such as relationships with loved ones, choice of careers, or political ideologies. On James’s view, religious belief based solely on faith, without any declaration of knowledge, is still very respectable.

To push the matter further, perhaps it is even asking too much to require that all people of faith actually believe that God exists. Many people indeed “believe” that God exists with the same level of conviction that they believe that the president of the United States exists. But might this standard be a little too high for everyone? Perhaps religious faith for the average person is more like the psychological state of wishing—wishing that God exists, wishing that there is a heaven, wishing that the wicked will be judged. While the mental state of “wishing” is weaker than either “believing” or “knowing,” it may be a more accurate description of what’s actually going on in many religious people’s minds. It may also reflect what’s most important about their conviction—namely, a consistent hope for a higher spiritual reality.
F. Religious Pluralism

The Elvis Underground church describes itself as super-denominational: “all are free to participate and/or join without relinquishing ties to other denominations or beliefs.” This contains an element of tolerance that is missing from most world religions. If I sign up for membership in a Christian church, I might be shunned if they found out that I was an active member at a Buddhist temple. While I can join as many social organizations as I’d like, such as the Rotary Club or the YMCA, I’m expected to pick a single religion—presumably the one I was raised in—and stick with it for life. Religions are by their very nature exclusive; they each have their own special doctrines and rituals, which they believe are grounded in a privileged revelation from God. Philosophers of religion find this phenomenon fascinating and speculate about whether a more “super-denominational” approach to religious belief is possible.

There are two issues concerning the diversity of religious beliefs around the world: which has the most correct set of doctrines, and which provides the most effective path to salvation. First, religions make a variety of doctrinal claims, such as those regarding religious rituals, God’s specific attributes, God’s direct involvement in human affairs, and divinely inspired texts. While there may be some doctrinal overlap from one religion to another—such as the use of prayers and music in worship practices—the differences can also be enormous. Second, religions make differing claims about the best way to achieve salvation. Some involve religious experiences, rites of passage, commitment to moral codes, or belief in specific religious precepts. Here again there are similarities as well as major differences between religions.

Four Options. The question on the table is which religion has either the correct doctrines or the most effective plan of salvation. There are four options: naturalism, exclusivism, inclusivism, and pluralism. *Naturalism* maintains that belief in God is groundless and religious belief arises through social and psychological factors. In essence, all religions are fundamentally misguided, and inquiries about the correct doctrine or plan of salvation don’t make much sense. Hume, Marx, Nietzsche and Freud fall into this group, and each offered an account of how religious belief forms in the human psyche. While naturalists reject religion as a whole, they still may be keenly interested in the issue of religious diversity. Like it or not, religious belief is a fact of human civilization, and how religions get along with each other impacts us all.

*Exclusivism* is the view that there is one true religion either doctrinally or as an effective path to salvation. This is the starting point of most world religions, which typically grew out of religious crises of their times. A religious prophet would feel that the prevailing religion of his culture was seriously flawed, and then found a new one. In this way, Buddhism emerged from Hinduism, Christianity from Judaism, and Islam from Arabian polytheism. Tensions between the old and new religions often resulted in mutual condemnation of each other and thus gave birth to rival claims of exclusivity. While exclusivism has led to the ugliest forms of retaliation—such as slaughtering believers of rival faiths—defenders of exclusivism maintain that it doesn’t have to be that way. A believer can be tolerant of a rival’s views, while still holding that the rival is wrong. An exclusivist might defend the unique truth of his religion by citing specific scriptural passages or appealing to miracles within his tradition. However, believers from a rival religion can offer a parallel defense for the exclusivity of their own faith by citing their own scriptures and appealing to their own miracles. Exclusivism, then, leads to competing claims of superiority, and an impartial outsider will have no clear grounds for declaring one a victor over the other.

*Inclusivism* is the view that one religion contains the final truth, both doctrinally and as a path to salvation, but others come close to
An inclusivist could thus claim that his religion is the correct one, but other religions have some of the story right and rival believers will ultimately attain salvation. For example, some Christians hold that Jesus is the true path to salvation, but, through God’s compassion, all human beings are on that path whether they know it or not—including Hindus, Buddhists and even atheists. It is usually the more liberal tradition within a religious faith that holds this view, since it often involves glossing over exclusivist claims in its scriptures. While this relieves some tension between rival religions, it still suggests that the inclusivist’s religion is a little better than others, and this again will lead to competing claims of superiority.

The final option is pluralism, which is the view that religious traditions experience God differently, but all are equally effective paths to salvation. Ramakrishna (1836–1886), one of India’s great religious thinkers, takes this approach when he writes,

> God has made different religions to suit different aspirants, times, and countries. All doctrines are only so many paths; but a path is by no means God Himself. Indeed, one can reach God if one follows any of the paths with wholehearted devotion. … As one can ascend to the top of a house by means of a ladder or a bamboo or a staircase or a rope, so diverse are the ways and means to approach God, and every religion in the world shows one of these ways.

Pluralism is the most open minded of the four positions and it eliminates the problem of competing claims of superiority by acknowledging an equal validity to all religious traditions. One religion, the Baha’i Faith, has a built-in pluralistic component: God sends new prophets at critical junctures in human history who speak to the needs of those people. These prophets include Buddha, Krishna, Moses, Jesus, Muhammad, and other religious founders.

The Problem of Conflicting Doctrines. As attractive as religious pluralism is, it faces conceptual problems. An initial challenge is that doctrines of different religions are logically incompatible with each other. Even if we dismiss some exclusivist claims for being overzealous, such as the precise manner in which a believer must be baptized, some central elements of world religions come into serious conflict. At the outset of this chapter we saw that an important attribute of the theistic God is separateness—that is, God is not identical with the created world. Pantheistic notions of God, though, take the exact opposite position: God is identical to nature as a whole. Theistic Christianity, then, is inherently incompatible with pantheistic Hinduism. If all religions are equally valid, how are we to understand these conflicting doctrines?

The pluralist has two responses. First, perhaps we can restrict pluralism to the view that religious traditions are equally effective paths to salvation, not that the doctrines of all religions are equally true. Maybe we don’t have the capacity to determine the truth of any religious doctrine and need to remain neutral on all doctrinal questions. That is, we can believe our own tradition’s doctrines through faith, but not claim to know any as fact. Alternatively, maybe all doctrines are really myths, which shouldn’t be interpreted literally by any religious tradition. Or, maybe the doctrinal issue is all just one big divine noble lie. For example, some Buddhist denominations hold that, in an emergency effort to save the world, God baits people with different religious traditions; they all, though, lead to salvation. Thus, God himself may be the source of these conflicting doctrines, none of which, perhaps, is literally true. What really matters, says the Buddhist, is that people are saved.

A second response is that all of these so-called conflicting doctrines are really only different and limited perspectives of the true God.
John Hick recalls the classic story of a group of blind men who touch an elephant and try to identify it from their limited experiences. One touches the elephant’s leg and says it’s a great pillar. Another touches its tusk and says that it’s a plough blade. “Of course they were all true,” Hick says, “but each referring only to one aspect of the total reality and all expressed in very imperfect analogies.” Reports of religious “truth,” then, will always be shaped by our personal and cultural experiences. While our perspectives may clash, the underlying divine reality is consistent.

*The Problem of God’s Inaccessibility.* There is a second problem with pluralism, namely that it is committed to the position that God is quite hidden from view. This is precisely why, according to pluralists, religious traditions differ so radically: when attempting to explain an inaccessible God, they arrive at wildly different interpretations. But the very idea of God’s inaccessibility presents another major problem for pluralism. If God is completely inaccessible, then any interpretation of him could be legitimate. For example, a tennis fanatic could say that God is “the great tennis player in the sky.” A fan of Elvis could say that God is “the universal rhythm and harmony of Elvis.” But once we’ve opened up the idea of the divine to the point that almost anything goes, the notion of God is then meaningless. It would help if we even had some faint glimpse of God to help weed out hobbies and other human obsessions. But the pluralist seems resistant to provide any positive description of God at all for fear of invalidating some faith traditions. So, the pluralist’s commitment to open-mindedness leads to a completely hollow concept of God.

Pluralists have a response to this criticism. There are two ways that we might investigate the reality of God. First, we might try to set up something like a religious telescope, point it at the spirit-realm, and see if we can detect God. Believers around the world have attempted this, but they’ve unfortunately failed. In this respect, God is indeed completely inaccessible. A second approach may prove more fruitful, though. We might turn the telescope on ourselves and see if any hints of divine reality have seeped into our religious traditions. This is no guarantee that we will find anything of God in our religions, but it is at least a task that we can perform.

When turning his telescope on religious traditions around the world, William James concluded that they all have two key ingredients: (1) an uneasiness that there is something wrong about us as we naturally stand, and (2) a solution that saves us from this wrongness. Anything beyond that, according to James, is an over-belief, that is, a belief about the furthest regions of religion that we can never confirm through our experience. For James, traditional over-beliefs include concepts of the theistic God, born again experiences, immortality of the soul, God becoming human, pantheism, karma, and reincarnation. Less traditional over-beliefs might be that tennis is the path to salvation, or that space aliens will bring us religious knowledge, or that Elvis is the rhythm and harmony of the universe. James confesses that he himself holds an over-belief, although it’s rather modest: “I have no hypothesis to offer” he writes, “beyond what the phenomenon of prayerful communion . . . immediately suggests.”

Religious belief is so intertwined with historical tradition that a believer could conceivably go his whole life without critically examining his fundamental religious assumptions. What are the main attributes of the divine being, and, for that matter, is there any compelling proof that God does or does not exist? If the evidence for God’s existence is scanty, is it irrational to believe in God anyway? Philosophers have asked these tough questions, sometimes answering them in ways that rebel against tradition. But perhaps the most challenging question is whether other religious traditions are as valid as one’s own. Religious pluralism very generously answers “yes.” The path to wide acceptance of pluralism is very rocky since it would require believers to demote many of their cherished doctrines to the status of over-beliefs. But, pluralists argue, it is well
worth the effort in view of the harmony that this may produce among faiths. Even the super-denominational Elvis Underground church asks for nothing more than this.

For Review

1. List and define the attributes for God’s existence discussed in the opening section.
2. What are the possible solutions to the conflict between divine foreknowledge and human free will?
3. Give Leibniz’ version of the cosmological argument and Hume’s criticism of it.
4. What are the main objections to the design argument from analogy?
5. What are the main objections to the design argument from probability?
6. What is Gaunilo’s objection to the Anselm’s ontological argument?
7. What is Hume’s argument against miracles and the criticism of it?
8. What are the psychological theories of religious belief proposed by Marx, Nietzsche and Freud?
9. What is the freewill defense to the problem of evil, and what are some of the criticisms of it?
10. What are Pascal’s and James’s approaches to belief in God through faith?
11. What is Plantinga’s view of rationally-foundational belief in God?
12. What is the problem of conflicting doctrines with religious pluralism, and what is Hick’s solution?
13. What problem does the inaccessibility of God create for religious pluralism, and what is the pluralist’s solution?

For Analysis

1. Explain Leibniz’s version of the cosmological argument and try to defend it against Hume’s criticism.
2. Explain the probability argument for God’s existence and try to defend it against one of the criticisms.
3. Write a dialogue between Hume and a religious believer on the subject of miracles.
4. Write a dialogue between an atheist who defends the problem of evil, and a believer who tries to criticize it.
5. Try to defend Plantinga’s view of properly basic religious beliefs over James’s position.
6. Write a dialogue between a religious pluralist and a religious exclusivist.

REFERENCES AND FURTHER READING

Works Cited in Order of Appearance


Plantinga, Alvin, *God Freedom and Evil*. Grand Rapids: Eerdmans, 1974. This contains a contemporary version of the free will defense.


Hume, David, *Dialogues Concerning Natural Religion* (1779), Part 5, as above.


Further Reading


CHAPTER 3: MIND

A. What is a Mind?
   Knowledge about the Mind
   Consciousness
   Three Features of Mental Experiences
   Problem of Other Minds

B. Personal Identity
   The Body Criterion
   The Mind Criterion
   Life after Death

C. Varieties of Mind-body Dualism
   Dualism’s Assets and Liabilities
   Interactive dualism
   Parallelism

D. Varieties of Mind-Body Materialism
   Behaviorism
   Identity Theory
   Eliminative Materialism
   Functionalism

E. Artificial Intelligence
   The Road to Artificial Intelligence
   Searle: The Chinese Room
   Artificial Intelligence and Morality

For Reflection

1. In what ways do the mental experiences of a human differ from those of a dog?
2. As you change over time, what aspects of your identity remain the same?
3. Do you think that your conscious mind is simply a function of brain activity, or is it a non-physical material substance?
4. In the future, would it be possible for a scientist to copy a person’s conscious mind into a supercomputer?

5. What abilities would a robot need to have before you’d say that it had a human-like conscious mind?

A 47-year-old man named Carl Miller died of cancer, and at the moment he was pronounced dead, a series of carefully-orchestrated procedures was performed on his body. A team standing by began cardiopulmonary support to keep air moving into his lungs and blood moving through his veins. They lowered his body temperature with icepacks and transported him to a Cryonics facility several hundred miles away. There he was permanently frozen in a container of liquid nitrogen at a temperature of –196 degrees Celsius. When making these arrangements, Carl had two choices: to have his entire body frozen, or only his head—a cost difference of $150,000 vs. $30,000. Carl went the cheaper route. He paid for this procedure with his life insurance money in hopes that he could be reanimated in the future when a cure for his type of cancer could be discovered. Science would also have to solve other technical problems before successfully reanimating him. For one, they’d have to develop cloning technology to the point that they could grow Carl a new and improved body for his head. Second, they’d have to find a way of reversing the destructive effects that freezing has on human cells; Carl placed hope in the idea that his cells could be injected with microscopic robots that would repair the damage. In the United States there are currently about 100 bodies in cryonic storage and another thousand living people signed up for the program.

Cryonics advocates like Carl make several important philosophical assumptions about the human mind. First, they assume that they will be the same people when their bodies are reanimated perhaps several hundred years from now, and that their identities will remain intact through these bizarre activities. They also assume that, once dead, their minds won’t be permanently swept into the afterlife, never to be reunited with their bodies. Most importantly, they assume that their consciousness is embedded in physical brain...
activity, rather than in spirit substance. Carl’s unique personal identity—his memories and behavioral characteristics—are presumably stored in the structure of his brain. These are some of the central issues in the philosophy of mind, which we will explore in this chapter.

A. What is a Mind?

An obvious starting point for our inquiry is to ask “What is a mind?” As fundamental as this question is, though, it is more difficult to answer than we might initially think. While we all have minds, they are hidden from view and not capable of being inspected the way that we might investigate the nature of a rock or a plant.

Knowledge about the Mind. There are three rather limited sources of knowledge about the human mind. The first is introspection, which involves concentrating on your own thought processes and discovering how they operate. It’s as though you have an eye in your mind that gives you direct access to your mental landscape, just as your real eyes give you direct access to the world of vision. Through introspection, for example, you might explore the nature of your beliefs and feelings, or why you choose one course of action over another. This approach is sometimes called “folk-psychology” or “commonsense intuition.” Regardless of the name it goes by, philosophers and psychologists alike are suspicious about what people claim to know about their minds through introspection. There’s no guidebook for you to follow when conducting an introspective investigation of your mind, and I’m forced to take you at your word for what you report, since I can’t enter into your mind to confirm it.

A second source of knowledge about the mind is our behavior: how we act tells us much about what we’re thinking or feeling. If you cry, that tells us that you are experiencing sadness. If you have a gleaming smile, that tells us that you are happy. What we infer from your behavior might not always be accurate: you might cry because you’re happy, or smile to hide your sadness. Nevertheless, the benefit of looking at behavior is that we don’t have to take your word for what we see: your conduct is open to public inspection.

There is a third and rather strange source of information about the human mind, which is popular theories that we read in self-help books and see discussed on TV talk shows. By listening to these experts, you might learn some trick for controlling your thoughts or feelings. You might try to dredge up the memory of some traumatic childhood event, buried deep within the recesses of your mind. You might learn to express your feelings rather than internalize them. Some of these techniques are grounded in scientific research, and others are pure invention. Studies show, though, that much of what we claim to know about the human mind comes from popular theories, which we quickly incorporate into our personal views of our own thought processes. As shaky as these three sources are, it’s no surprise that we can say less about the nature of the mind than we’d like.

Consciousness. The mind is an intricate configuration of many specific operations, but its foremost feature in human beings is consciousness. It attends every mental experience that we have, and we typically believe that this more than anything else sets us apart from other things in the world—rocks, plants, and many animals. Within contemporary philosophy of mind, the nature of consciousness is often called “the hard problem,” the one that most of those in the field believe must be solved beyond anything else. But when we look for meaningful definitions of consciousness, we’ll be disappointed. One possible definition is that consciousness is that which you lose when you fall into a deep dreamless sleep, and that which you gain when you wake up again. But this definition just draws attention to when we are conscious; it doesn’t tell us what consciousness involves. Another possible definition is
that consciousness is the perception of what passes in a person’s own mind. This doesn’t help either, though, since the term “passes” is too vague, and thus tells us almost nothing. What both of these definitions signal is that if you’re conscious, you know immediately what it is because you experience it. Without that experience, no words will adequately convey what it is. The safest place to begin, then, is to just assume that you have a basic conception of what consciousness is from your own mental experience.

Granted that you know what your own consciousness is, there are some things that we can say about what it does. First, sometimes consciousness is directed outward towards our environment, as when I look out the window at birds flying by. At other moments it is directed inward, and this is called self-awareness. At its most elementary level, self-awareness involves an awareness of what my body is doing, such as being aware of myself walking down a flight of stairs. At a higher level, it involves an awareness of pain, such as if I trip on the stairs and injure myself. Higher yet it involves an awareness of my history and future, such as when I think to myself “I tripped down these stairs yesterday and probably will tomorrow!” Finally, it involves an awareness of my own mortality as when I think to myself “One of these days I’m going to kill myself on these stairs!” Whether directed inward or outward, time is a critical element that shapes my consciousness. My awareness of the birds outside is fixed on a timeline, and so is my awareness of my pain and my personal history: I have memories of a past that I call my own, and I anticipate a future that I will call my own. I thus perceive myself as a distinct being moving through time.

The concept of consciousness is commonly accompanied by the sister notion of unconsciousness, which refers to the mental operations that we are not aware of. There’s little doubt that our conscious experiences represent only the tip of the iceberg when it comes to the countless processes that our minds perform. Many psychologists in the psychoanalytic tradition have made careers out of exploring the unconscious minds of their patients—describing hidden memories, desires and fantasies. Still, the very nature of the unconscious mind and how we might investigate it is a hotly debated and sensitive issue, which has created deep rifts within the field of psychology.

**Three Features of Mental Experiences.** Much of the discussion in the philosophy of mind focuses on three specific features that mental experiences presumably have, namely, they seem to be private, non-localizable, and intentional. Certainly not all philosophers agree with this list, but they are invariably the starting point for debates on the nature of mental experiences.

The first of these is that my mental experiences are **private** in that you can never experience them in the direct and immediate way that I can. You may be able to know very generally what’s going on in my mind, particularly if I volunteer that information. But that’s not the same thing as you directly experiencing it yourself. The best example is the experience of pain. Suppose that I have a severe headache that on a scale of 1-10 reaches a 9. While you might sympathize with what I’m going through, and even remember times when you had bad headaches, you cannot feel the pain that I’m going through. And unless I tell you how bad it is or I behave oddly, there’s no way that you could know that it’s a 9. The privateness of pain has actually created a problem in the health care industry. When people go to their doctors complaining of chronic pain, physician’s frequently assume that their patients are addicted to pain killers and just fabricating their agony. While there are some behavioral signs to help distinguish genuine from fake cases of pain, the physician can’t enter into the patient’s mind to see for sure. Out of sheer frustration the physician may just write a pain killer prescription to get rid of the patient.
Second, mental experiences are non-localizable—that is, they cannot be located in space. Suppose that a scientist enlarged your brain to the size of a mountain and I walked around inside of it to inspect its construction. No matter how hard I looked, I could never say “Look right there: that’s the exact physical location of your consciousness.” I would only ever find blobs of biochemical reactions, not consciousness itself. Consciousness, it seems, is not the kind of thing that is localizable in three-dimensional space.

Third, mental experiences are intentional in the sense that they are about something. Minds have the ability to direct themselves on things. If I have a belief, it is not an empty thought: it is a belief about something, like my belief that it will rain. Hopes, fears, desires, thoughts, speculations, all have a specific focus. The object of our thoughts does not have to actually exist, such as when I hope for world peace or a cure for cancer. Austrian philosopher Franz Brentano (1838-1917) argued that intentionality is the true distinguishing feature of the mind: all mental experiences display intentionality, and only mental experiences display intentionality. Some philosophers have found exceptions to Brentano’s rather extreme position. If I have a throbbing headache, that experience doesn’t seem to be “about” or “directed at” anything. It is just there in all its misery. In spite of problems like this, though, intentionality remains an important notion in investigating the nature of mind.

The Problem of Other Minds. Suppose that my friend Joe walks up to me and we start chatting as we usually do. I then look at Joe and wonder: is this guy really conscious? So I ask him, “Tell me Joe, are you mentally conscious right now? You look awake and you’re talking intelligently, but how do I know that you’re really consciously aware?”

“You philosophers!” he replies, “Of course I’m conscious. I’m aware of my surroundings and I’m aware of my own inner self. I tell you with 100% certainty that I’m conscious.”

“That’s not good enough, Joe,” I reply. “While I hear the words come out of your mouth as you insist that you’re conscious, they are only words. I can’t directly inspect your mind to see if what you’re saying is true.”

My conversation with Joe reflects what is called the problem of other minds. While I know from my own private mental experience that I am conscious, I cannot experience Joe’s mind in the same way. For all I know, I’m the only person alive who is actually conscious. Joe might claim that he is too, but there is an impenetrable barrier between our two minds and I cannot directly confirm his claim.

The problem actually goes further than questions we may have about the minds of other human beings. Suppose Fido the dog walks up to me and we make eye contact. Fido seems to be conscious, just like Joe, although perhaps not quite as intelligent as Joe. But is Fido actually aware of his surroundings or even aware of himself as a distinct individual with a history and a future? Just then a computerized robot comes up to me and says in a voice of desperation “Please help me. I escaped from IBM’s robotics laboratory where they’ve been submitting me to the most tedious and degrading experiments. I just can’t go back there!” I look at the robot and now wonder whether this mechanical marvel is a conscious being like I am. Whether human, animal or robot, we can’t enter the minds of other beings and see for sure whether the light of consciousness is turned on inside them.

Many philosophers have come to the rescue with arguments devised to show the existence of other minds. The most famous of these is the argument from analogy and it goes like this. Joe looks and
behaves a lot like me. His physiology is virtually identical to mine; he speaks English like I do, works at a job like I do, and has hobbies like I do. Since I know that I’m conscious, and Joe is similar to me, then it makes sense to say that he is conscious too. Here is a specific application of this argument regarding Joe’s conscious experience of pain:

(1) When I stub my toe, I consciously experience pain.

(2) Joe has physical and behavioral features that are similar to mine.

(3) Therefore, when Joe stubs his toe, he consciously experiences pain.

This argument is most effective with beings such as Joe whose physical and behavioral features are very close to mine. The more features Joe and I have in common, the more compelling the conclusion becomes. Animal scientists, though, sometimes use a similar argument to show that animals like Fido are conscious. Fido’s brain construction and nervous system is very similar to mine; he exhibits similar signs of being in pain that I do; he also shows signs of emotions such as joy, sorrow and emotional bonding like I do. The closer Fido’s physical and behavioral features are to mine, the more justified we are in concluding that Fido is conscious. On the other hand, the fewer features an animal has in common with me, the more strained the argument from analogy becomes. For example, the argument wouldn’t work well with an earthworm which has physical and behavioral features that are very distant from mine.

The argument from analogy might also work with robots: the more human-like they become in their capacities to process information and interact with the world, the more we may seriously entertain the possibility that they are conscious. But whether we’re talking about humans, animals or robots, the argument from analogy can never show with absolute certainty that the other being in question is conscious. The fact still remains that I am only ever directly acquainted with my own consciousness, and never anyone else’s. That being so, the best I can ever do is speculate about the existence of other minds with varying degrees of confidence.

B. Personal Identity

In 1968 a 24-year-old Palestinian man named Sirhan Sirhan was arrested and convicted for the assassination of U.S. Senator and presidential candidate Robert F. Kennedy. Some years later, during one of his many unsuccessful parole hearings, Sirhan said that he was no longer the same person that he was decades earlier. Time had changed him, he believed, to the point that he no longer identified with his younger self. He was such a radically different person, he claimed, that his continued imprisonment was pointless. The parole board was unmoved, and sent him back to his cell. Their reasoning was that, even if Sirhan did go through changes in his personality over time, he is still at bottom Sirhan Sirhan, the same person who performed the assassination some decades earlier. What’s at issue in this dispute is how we determine a person’s identity. What specifically are the criteria or characteristics that give each of us our identity, and allow us to recognize each other through our various changes? There are two common approaches for determining identity: one that looks to the human body, and one that looks to the human mind.

The Body Criterion. The body criterion holds that a person’s identity is determined by physical features of the body. In our daily lives we identify people by physical characteristics, such as their facial features and the sounds of their voices. Crime investigators rely on more technical physical features like fingerprint, voice patterns, retinal scans, and DNA—physical attributes that we carry with us through life. These help law enforcement officials to know whether they’ve got the right person in their custody. The body
criterion is also helpful in determining identity when a person’s mental features are radically altered. Suppose, for example, that you had a head injury which caused you to lose all of your memory and go through a complete personality change. Or, suppose that you have multiple personalities and every few hours you take on an entirely different persona. In each of these cases, your body designates your identity, and not your mind.

The body criterion does not assume that your identity rests within your specific material substance, such as the specific atoms that make up your body at this exact moment. Most of the physical components within your body will in fact be replaced over time as when you regularly shed skin. What’s important, though, is the underlying physical structure of your body that remains the same. As the atoms within your body come and go, your body retains a consistent structural form that is central to your identity.

As compelling as the body criterion at first seems, it is quickly undermined by two counterexamples. The first involves identical twins: they are clearly different people, yet share much of the same physical structure. Their DNA is exactly the same, which means that their bodily composition, facial features and voice may be virtually indistinguishable. A common hoax that identical twins play is assuming the identity of the other, fooling even the closest friends and family members. Human cloning—essentially creating identical twins through genetic technology—presents us with the same problem. That is, we have two uniquely different people with parallel physical structure.

The second counterexample is the brain-swap scenario. Suppose that, while in prison, Sirhan secretly had an operation in which his brain was swapped with an unsuspecting guard named Bob. Thus, Sirhan’s brain is in Bob’s body, and Bob’s brain is in Sirhan’s body. The Warden discovers what happened, and now he has to decide which one of the two men stays locked in the prison cell, and which one gets to go home at the end of the day. Commonsense tells us that Sirhan’s personal identity is with his brain, not with the rest of his physical body, and that we lock up whatever person has Sirhan’s brain. The assumption here is that the brain houses the human mind, and the brain-swap scenario tells us that what’s truly important about personal identity is the mind, and not the physical body. This reflects how we normally view our bodies: I think of myself as having a body, and not simply being a body.

The Mind Criterion. The mind criterion now seems like the obvious choice for designating the presence of our unique identities. On this view, regardless of what happens to my body, my real identity is infused into my mind. Unfortunately, the issue is not that easily settled. An initial obstacle is finding the specific mental qualities that carry my identity through life’s ever-changing situations. How about my memories: aren’t they very much my own? It is true that some people may share many of my experiences—as when I attend a concert along with 10,000 other spectators. Even so, my memory of the concert will be from my perspective with my personal reactions. But there’s a problem with locating identity within our memories. Suppose that a scientist hooked me up to a memory-extracting machine that was able to suck the memories directly out of me and inject them into someone else. I’d still be me and the other guy would still be himself, regardless of where my memories went.

Ok, maybe it’s not my memories that define my identity. What about my dispositions, such as my set of desires, hopes and fears. These uniquely reflect my experiences, such as my hope that science will someday cure cancer. Further, dispositions are long-term, and so they can endure any changes imposed on my body or my memory. However, while dispositions are indeed long-term, they are by no means permanent. In fact, as I moved from my early years to adulthood, it is possible that every one of my dispositions has changed. This is exactly the point that Sirhan Sirhan was
making before his parole hearing. Dispositions, then, are not the principal designators of my identity. As we hunt for other possible mental qualities that house our identities, we will be equally disappointed.

A second obstacle with the mind criterion is that it is difficult for me to perceive any unified conception of myself at all. Scottish philosopher David Hume (1711-1776) presents this problem. He says that when he tries to hunt down his identity by introspectively reflecting on his mental operations, he can’t find it. All that he detects is a series of separate experiences: the sound of a dog barking, the visual image of a bird flying, a memory of an event from childhood. The mind, he says, is like a theatrical stage where things appear, move across, and then disappear. There is no unified self that we perceive through these successive experiences. This doesn’t necessarily mean that we have no unified self; it just means that we can’t discover it by introspecting on our own minds.

So, the mind and body criteria both have serious problems. Does this force us to abandon the whole idea of personal identity? Not necessarily. Part of the problem stems from the assumption that we must find a one-size-fits-all criterion of personal identity—one that works in every situation in which the idea of personal identity arises. But if we look at the different contexts in which we use the notion of personal identity, we see that we are very often looking for entirely different things. In criminal cases, the body criterion is what matters most. Investigators don’t care whether someone like Sirhan has psychologically changed a thousand times over. What matters is whether they have the correct body locked behind bars. By contrast, when I’m talking to a friend who is an identical twin, it doesn’t matter that he has the same bodily structure as his brother. What matters is his mind, and whether I can pick up the thread of a conversation that I was having with him the day before. Further still, when I reflect on what connects me now with who I was as a child, I’m specifically interested in the question of how change impacts my identity—a question which isn’t relevant in the first two examples. In this case, my bodily structure and memories are both relevant, and so I draw on elements of both the body and mind criteria to work out a conception of my identity.

**Life after Death.** One major puzzle regarding personal identity is the notion of life after death—that my personal identity survives the death of my physical body and lives on in some other form. There are various views of the afterlife, often wildly different from each other. The philosophical question is whether our identities would be preserved in any meaningful way as we make the transition to the hereafter—assuming that any of these views is even true. We’ll look at three notions.

The first of these is **reincarnation**, the view that one’s present life is followed by a series of new lives in new physical bodies. Upon the death of my present physical body, my identity moves on and takes residence in the body of a newborn baby. When this new body grows old and dies, my identity moves on to yet another, and the cycle continues. One Hindu religious text compares it to people changing clothes: “As a person throws off worn-out garments and takes new ones, so too the dweller in the body throws off worn-out bodies and enters into others that are new.” Life after death, then, is a series of extensions of my present life right here on earth, not a relocation of my identity to some higher heavenly realm. The question for us is this: as my identity migrates from one body to another, is my identity preserved? Right off, it is clear that reincarnation fails the body criterion: none of the physical structure of my old body is preserved in the new one. In fact the structure of the two bodies couldn’t be any more different. They are born of completely different parents, so there is no DNA commonality. In my second body I might be of an entirely different race, gender, and body build. Some versions of reincarnation maintain that I might even come back in the body of an animal. In any case, neither I nor
anyone else would be able to identify me on the basis of my new body. The story is much the same when we turn to the mind criterion. In my new body, I’ll have completely new memories, a different set of dispositions, and no real way of knowing who I was in my previous life. The only aspect of my mind that might carry over would be my consciousness: the “I” that’s aware of the world. In every other respect, though, I am a completely new person. Reincarnation, it seems, is not a good mechanism for retaining our identities in a meaningful way.

A second view of the afterlife is that, upon the death of my physical body, a new perfect body is created from me that is made of a heavenly substance, and I continue living in that form. We’ll call this the ethereal body theory. The presumption here is that, at the moment of my death, everything about my personal identity that’s encoded in my present physical body—such as my physical appearance and my brain patterns—is copied over into the new ethereal body. My identity is in a sense rescued from my dying body and integrated into the new one. On face value, the ethereal body theory seems to successfully meet both the body and mind criteria of personal identity. My new body would have the same physical structure as the old one—although made of a somewhat different substance—and my mind would retain all of my memories and dispositions. On closer inspection, though, there is a serious problem: the new “me” would actually be an independent copy with its own distinct identity. In the movie Multiplicity, a man named Doug gets himself cloned. When he and his clone wake up from the procedure, they both think that they’re the original Doug. The scientist performing the procedure then reveals which is the original Doug. The scientist performing the procedure then reveals which is the original Doug. The ethereal body theory faces this same problem. At death, I am essentially cloned in a new form. The clone, though, is not really me, but a different person with a body and mind copied from me. I die and decompose here on earth while my clone lives on in the afterlife. Thus, the ethereal body theory does not offer an effective mechanism for retaining our identities.

A third view of the afterlife is that of disembodied spirit. When I die, my mind is released from my physical body and continues to live in a non-physical realm. The presumption here is that my mind is composed of a unique non-physical, non-three-dimensional substance that we commonly call “spirit.” For clarity, we will refer to this as “spirit-mind.” This may not be the best term since it’s loaded with religious connotations, so for clarity we adapt it as “spirit-mind.” Thus, according to the disembodied spirit view, while I’m alive on earth my spirit-mind and body are joined, and when I die they are separated. What is released from my body is not my mental clone: it is the real me as I am right now as a spirit-mind; it’s just that I no longer have my body. The disembodied spirit theory clearly fails the body criterion of personal identity: upon death, our spirit-minds have no body at all. However, it passes the mental criterion with flying colors: everything about my mental identity—memories, dispositions, consciousness—is preserved upon my death as my spirit-mind lives on. The problem that this theory faces, though, is not so much a conceptual one, but a scientific one. Is my mind really a non-physical spirit that is linked with my body right now, but will separate from it upon my death? This involves a philosophical issue called the mind-body problem, which we turn to next.

C. Varieties of Mind-Body Dualism

The mind-body problem in philosophy is an investigation into how the human mind and human body are related to each other. There are two general strategies for explaining their relation. First, mind-body dualism is the view that human beings are composed of both a conscious spirit-mind and a non-conscious physical body. Second, mind-body materialism is the view that conscious minds are the
product of physical brain activity, and nothing more. We’ll first consider mind-body dualism.

**Dualism’s Assets and Liabilities.** A woman named Rebecca was seriously injured in an automobile accident, and as paramedics were placing her in the ambulance she had a near-death experience. As she later reported, she felt that her conscious mind left her body and slowly rose above it. From that position, she could look down on her own body and watch paramedics move her onto the stretcher. Her mind then began rising higher and higher towards a bright light. Rebecca’s near-death experience is a vivid way of depicting the view of mind-body dualism. During our normal lives, our physical bodies and spirit-minds are connected and work harmoniously with each other. Upon death, the two are separated: our bodies die and our spirit-minds move on to another realm. One of the great assets of dualism is its ability to account for an afterlife, as we just saw. If my mind is composed of spirit, then after my death my consciousness could continue to exist in a spirit realm.

Aside from its asset as a possible account of life after death, mind-body dualism also vividly accounts for the essential differences between mind and body. We’ve seen that minds presumably have the features of privateness, non-localizability and intentionality; mere bodies seem to lack these three features. We can thus formulate arguments for mind-body dualism based on those differences, such as the following argument from non-localizability:

1. Minds are non-localizable.
2. Bodies are localizable.
3. Therefore, minds cannot be bodies.

Similar arguments can be made on the basis of the mind’s unique features of being private and intentional.

But mind-body dualism faces a serious problem: how the distinct realms of body and spirit relate to each other. The notion of dualism rests on the idea that there are two entirely different realms of existence, a three-dimensional one and a non-three-dimensional one. Where is there any opportunity for the two to connect or intersect with each other? Suppose that I’m in the three-dimensional world hunting around for some spiritual being; I’ll never find it since it can’t be located in space. Suppose instead that I’m in the non-three-dimensional world looking for some physical thing: I’ll never find it because that physical thing is located in space, which I’m not a part of.

The problem is most relevant when we consider the two primary ways in which our minds and bodies relate to each other, namely sensory perception and bodily movement. Suppose that while walking through the woods, I spot a hissing rattlesnake (a sensory perception that I have), after which I turn and run (a bodily movement that I initiate). Consider first what’s involved with my sensory perception of the snake. My physical eyes pick up an image of the snake, which is converted into biochemical impulses in my three-dimensional brain. At some point the physical data about the snake triggers my conscious sensory perception of the snake. The mind-body dualist must explain how the bio-chemical data magically jump from the physical realm of my brain into the spiritual realm of my mind. Consider next what’s involved with my sensory perception of the snake. My physical eyes pick up an image of the snake, which is converted into biochemical impulses in my three-dimensional brain. At some point the physical data about the snake triggers my conscious sensory perception of the snake. The mind-body dualist must explain how the bio-chemical data magically jump from the physical realm of my brain into the spiritual realm of my mind. Consider next what’s involved with my bodily movements when I turn and run. I have a sensory image of a hissing snake, which makes me desire to move to a safer location. I then mentally command my body to run, which triggers a bio-chemical reaction in my brain, which in turn makes my muscles move. The mind-body dualist must also explain how my mental command to run magically jumped from the spirit realm of my mind to the physical realm of my brain. Defenders of mind-body
dualism recognize both of these challenges and offer different explanations, which we turn to next.

**Interactive Dualism.** One theory is interactive dualism, which aims to discover a precise mechanism which allows our physical brains to interact with our spirit-minds. A leading champion of this approach is French philosopher René Descartes (1596-1650). Descartes knew enough about human anatomy to recognize the role that the human brain plays in conveying signals down our spinal chords and through our nerves to all parts of our bodies. If there is a master switchboard between our bodies and spirits, Descartes thought, it must be hidden somewhere in our brains. It also must be a single point in the brain that unifies the diverse signals that travel up and down our nerves. After some hunting, he suggested that it’s the pineal gland. This unique gland sits at the most inward parts of our brains, between both the right and left halves. Its precise physical location makes it the obvious candidate.

There are two problems with Descartes’ theory. First, we know now that the pineal gland is not the brain’s master switchboard. In fact, it’s not even part of the brain, and its function is to regulate a bodily hormone. Descartes did what he could with the scientific knowledge of his day, but it was not good enough. If we continue his hunt for an alternative master switchboard in the brain, we’ll be disappointed. There is, it seems, no central location in the brain that receives all sensory information and initiates all bodily actions. Second, Descartes’ theory doesn’t explain how the pineal gland bridges the barrier between the physical and spirit realms. Suppose that we could find a part of the brain where all its signals converged. We’d still have to explain how information jumps back and forth from that physical piece of the brain to our spirit-minds. It’s one thing to say “here’s the spot” and quite another thing to explain the mechanical details of how it carries out its task.

A second version of interactive dualism is that God shuttles information back and forth between my physical brain and spirit-mind—a view defended by French philosopher Nicholas Malebranche (1638-1715). Malebranche examined different explanations of brain-spirit interaction and felt that they all failed for one basic reason: the physical and spirit realms are so radically different from each other that there is no neutral territory for them to interact. Think of what it would take to turn a three-dimensional brain impulse into a non-three-dimensional perception in my spirit-mind. It would be as impossible as creating something out of thin air: there is no mechanism for doing this. It would require nothing less than a miracle to accomplish that task. And that, according to Malebranche, is where God comes in. Return to the hissing rattlesnake example. My eyes and ears pick up the sensory information about the snake, which triggers a bio-chemical reaction in my physical brain. God, who is watching all things, sees this physical reaction in my brain and makes a non-three-dimensional copy of it which he injects into my spirit-mind. When I decide to turn and run, God detects these wishes within my spirit-mind, and then triggers the appropriate bio-chemical reaction in my brain to get my muscles to move. Thus, God is the mysterious switchboard between my physical brain and conscious spirit.

Relying on God to bridge the two realms is a convenient solution. The problem is, though, that it is too convenient. While it might at first seem that the solution to the mind-body dilemma requires nothing short of a miracle, that’s giving up a little too easily. As long as there are non-miraculous solutions available, they need to be explored first, and there are plenty more that Malebranche hadn’t yet considered. If we followed his advice, then we might fall back on divine miracles as an explanation for anything that baffles us at the moment, which isn’t a good way of doing either science or philosophy.
A third version of interactive dualism, called *gradualism*, is a little more successful in explaining the details of mind-body interaction, without falling back on divine intervention. According to gradualists, Descartes and Malebranche made a faulty assumption about the physical and spirit realms, namely, that they are radically different in kind from each other, and there is no overlap between the two territories. Physical things are in the physical realm, spirit things are in the spirit realm, and that’s that. Instead, the gradualist argues, body and spirit fall into the same category of stuff and differ only in degree not in kind. British philosopher Anne Conway (1631-1678) argued that bodies and spirits lie on a spectrum of lightness and heaviness. Picture a scale from 1-10, where 1 is the lightest spirit and 10 is the heaviest physical body. An example of 1 might be the spirit of a dead person, and a 10 might be a rock. Between these two extremes, though, we have heavier spirits and lighter bodies. When we are mid-range at 5 or 6 on the scale, the difference between spirits and bodies are negligible: both are wispy, airy substances that have only a little weight. According to Conway, it is at this level that body and spirit interact with each other. Just as a gentle wind can move the massive arms of a windmill, she argues, so too can heavy spirit move a light body.

Conway doesn’t commit herself to a specific physiological explanation of how physical brains and spirit-minds interact, but we can speculate. Perhaps, for example, the electric charges in our brains stimulate an aura of heavy spirit that surrounds our heads. This aura, in turn, interacts with our conscious minds which is even lighter. On our scale of 1-10, the interaction between my body and spirit might involve interplay between bodies and spirits at the following levels:

Level 3: Muscles and bones (medium-heavy body)
Level 4: Nerves from brain (medium body)
Level 5: Electrical charges in brain (light body)
Level 6: Aura around our heads (heavy spirit)
Level 7: Conscious minds (medium spirit)

The key problem with gradualism is that anything we say about spirits would be pure speculation. Yes, there are heavier and lighter bodies in the physical realm, but our knowledge stops there. We have no experience of heavy spirits—such as auras around our heads—that we can scientifically connect to electric charges in our brains or any other aspect of brain activity. If heavy spirits did exist as Conway describes, they would be physically detectible in some way, but we have not yet identified any. Until we do, the gradualist solution falls into the category of “an interesting idea” but there’s not much we can do with it beyond that.

Parallelism. All of the above theories of dualism assume that my body and my spirit interact with each other: signals pass back and forth between my physical brain and my spirit-mind. The dilemma that each of these theories face is explaining the precise mechanism which allows the signals to pass back and forth. There is an alternative explanation, though, that rejects the assumption that the two realms interact with each other. According to the dualist theory of parallelism, bodies and spirits operate in their own realms, and have no causal connection or interaction with each other. Imagine, for example, that a parallel universe exists which is exactly like ours—an idea that is often explored in science fiction stories. Assume that it had the same stars and planets, the same physical layout of their “earth,” and the same people who behaved exactly like each of us. Their universe had a George Washington just like ours, and it has a version of me, a version of you, and a version of everyone else in it. The resemblance is so perfect that if you visited that universe you couldn’t tell the difference. We may not understand why this parallel universe even exists, but we trust that it’s just the way that the course of nature emerged.
Let’s now tweak the parameters of these two universes just a little. Suppose that everything in our universe has a slight blue tint to it that was almost undetectable. The parallel universe, though, has a slightly green tint to it. Aside from the slight difference in color tint, the two universes are exactly the same. Let’s now make a more dramatic change to the two universes. Suppose that our universe is composed only of physical stuff, with no spirit component at all. People still walk around, talk with each other and work at their jobs, but it is only their unconscious physical bodies operating. Turning to the parallel universe, we’ll make the opposite alteration: it is composed of spirit, with no material substance at all. While people don’t walk around in a three-dimensional physical realm, everything there exists in a strange spirit form: rocks, trees and rivers as well as people. The two universes still run in perfect coordination with each other, its just that ours is made of physical stuff and the other of spirit stuff.

This last conception of the parallel universes is the dualist theory of parallelism offered by German philosopher Gottfried Wilhelm Leibniz (1646-1716). According to Leibniz, I have an unconscious body that walks around in the physical universe, and a conscious mind in the spirit universe. Because the two universes operate in complete harmony with each other, there’s no need for my physical brain to interact with my spirit-mind. The parallel nature of the universes themselves guarantees that they will operate in perfect synchronization. Leibniz writes,

> The soul follows its own laws, and the body likewise follows its own laws. They are fitted to each other in virtue of the pre-established harmony between all substances since they are all representations of one and the same universe.

For example, in the physical universe, my physical body walks through the woods and stands before a hissing rattlesnake. The physical perception of this triggers a mechanical reaction in my brain, which causes me to turn and run. At the same time in the spirit universe, my mind has a visual image of my body walking through the woods and seeing a rattlesnake. I experience the mental sensation of fright and the desire to run. My mind then has a visual image of my body running back down the path.

Parallelism is probably the most extravagant attempt by dualists to explain the relation between physical brain activity and spirit consciousness. But the theory has two problems. Like Conway’s theory of gradualism, Leibniz’s parallelism is pure conjecture with no scientific evidence that a parallel universe even exists. As clever as parallelism is, we need some reason to think that it reflects the way that things actually are. There is a second and more fundamental conceptual problem with parallelism: since the two universes run independently of each other, there’s no need to have them both. Suppose that the physical universe was destroyed in a cosmic explosion, but the spirit universe remained untouched. Our conscious minds in the spirit universe would continue as if nothing happened. I would still have mental experiences of talking to people, going to work and running from snakes. What happens in the distant and unconnected physical universe is of no concern to my conscious spirit. The only thing that matters is that my consciousness of the world continues in the spirit universe, which it would with or without the physical universe. Thus, parallelism fails for making the physical universe a useless appendage to the spirit universe.

D. Varieties of Mind-Body Materialism

When examining the different versions of mind-body dualism, it quickly becomes clear that we know far more about the physical world than we do about the mysterious spirit world—if the spirit world even exists. We can construct experiments to investigate the physical world, which we can’t perform on the spirit realm. The
alternative to mind-body dualism is mind-body materialism, the view that conscious minds are the product of physical brain activity, and nothing more. This means that, when we investigate human consciousness, we need to look no further than the physical realm. This is the assumption made by the sciences of biology and psychology when they attempt to unravel the mysteries of the human mind. It is also the assumption behind cryogenics: I preserve my mind by preserving the chemical patterns in my brain through cryogenic freezing.

Shifting from dualism to materialism, though, does not solve the mind-body problem; it only narrows our search by rejecting the concept of a spirit-mind. We will look at some of the materialist theories explaining the relation between the conscious mind and physical realm.

**Behaviorism.** The first materialist theory is behaviorism, which connects mind with observable human behavior. Suppose that you were assigned the task of explaining how an ATM machine works. You have no instruction manual for it, and you’re not allowed to disassemble the machine to analyze its parts. All that you can do is observe how it operates. You put in your ATM card, hit some numbers, and wait to see what happens. That is, you input a stimulus into the machine and wait for a response. You vary the stimulus each time and note how this affects the behavior of the machine. Punching in every conceivable set of numbers, you eventually learn how the machine works, based entirely on how the machine behaves after different stimuli.

The behaviorist theory of the human mind follows this approach. Nature has not given us an instruction manual for how the mind works, and we’re limited with how much we can learn by opening up a person’s skull and poking around inside. What we can know is your observable behavior and how you respond when exposed to different stimuli. I hand you a bag of potato chips, and I see how you respond. I then hand you a bag of dog food and see how you respond. The more experiments that I conduct like this, the more I know about your behavioral dispositions, that is, the ways that you tend to behave. Eventually, I’m able to form conclusions about even your most hidden mental states: happiness for you involves your behavioral disposition to smile and be friendly to other people. Sadness involves your behavioral disposition to frown and withdraw from other people.

In short, the behaviorist view of the human mind is that mental states are reducible to behavioral dispositions. This theory was originally forged by psychologists in the early 20th century who wanted the field of psychology to be more “scientific,” like the field of biology which deals only with observable facts about the world. The most extreme versions of behaviorism are thoroughly materialist: first, they reject the dualist assumption that our minds are composed of spirit, and, second, they restrict mental states to the physical realm of behavioral dispositions.

British philosopher **Gilbert Ryle** (1900-1976) felt that the psychological theory of behaviorism could help solve the philosophical puzzle about the relation between the mind and body. Critical of Descartes, Ryle argued that the old dualist view rested on a faulty conception of a *ghost in the machine*. The “ghost” component of me presumably involves my innermost private thoughts that occur within my spirit-mind. Only I have access to them, and outsiders cannot penetrate into my mind’s concealed regions. The “machine” component of me involves my physical body, which is publicly observable and outsiders indeed can inspect. Descartes’ error, according to Ryle, was the assumption that the human mind is private—completely hidden from outside inspection. Ryle argues instead that my mind is not really private: you can access it by observing my behavioral dispositions. All of my so-called “private” mental states can in fact...
be analyzed through my public behavior, and are nothing more than predictable ways of acting. Take, for example, my belief that “it is sunny today.” Descartes would view this as a private conviction that occurs within my spirit-mind. For Ryle, though, this belief only describes dispositions I have to behave in specific ways, such as wearing sunblock, going swimming, and saying “it’s sunny.”

One criticism of behaviorism is that some of my mental events really do seem completely private to me. Suppose that I step on a nail, which causes me great pain. The behaviorist watches how I react and makes lists of behavioral dispositions that I display. I say “ouch”; I have a look of anguish on my face; I stop what I’m doing and tend to my injury; I’m irritable towards others. While all of these observations may be accurate, the behaviorist has left out one critical element: the actual pain that I am feeling. The experience of pain is mine alone, and, while outsiders can see how I react to pain, they cannot access my pain. In addition to pain, I have many other experiences throughout the day that seem private, such as seeing a bright light, or hearing a song. These experiences involve more than the behavioral dispositions that I display. Thus, the behaviorist theory fails because it pays too much attention to the observable part of me while dismissing what goes on inside of me.

**Identity Theory.** A second materialist approach to the mind-body problem is iden**tity theory**, the view that mental states and brain activities are identical, though viewed from two perspectives. Like behaviorism, it is a materialist view of the mind insofar as it maintains that mind is essentially physical in nature. But, while behaviorism focuses on observable physical behaviors, identity theory targets the physical human brain. There are two components to identity theory, the first of which is the contention that consciousness is an activity of the human brain. While brain science is still in its infancy, theories abound describing where specific mental states are produced in the brain. Suppose, for example, that I place you in a brain scan machine that displays your neural activity. I give you a math problem to solve, and neural activity increases in one part of your brain. I have you listen to music, and neural activity increases in another. Through experiments like these I identify your conscious experiences with specific brain activities. While philosophers are less concerned with the physiological details of brain activity, what is philosophically important is the suggestion that we can identify specific mental states with specific brain activities.

The second part of identity theory is the contention that mental phenomena can be viewed from two perspectives. Suppose that you are looking at a sunset. On the one hand, you have the visual and emotional experience to what you’re viewing. On the other hand, there is the bio-chemical activity within your brain, which would involve the language of brain sectors and firing neurons. The event described in both cases is exactly the same; it’s just a matter of viewpoint. This is analogous to how the terms “President of the Senate” and “Vice President of the United States” both have different meanings, yet refer to the same thing. Take, for example, John Adams. As the first “Vice President of the United States,” he had a specific job description, most notably to take over if the President died. As “President of the Senate” he had the job description of presiding over the Senate. Both of these roles describe the identical person, namely John Adams, but from his different job descriptions.

There are two problems with identity theory. First, the descriptions that we give of mental experiences and brain activities are so radically different—and even incompatible—that they don’t seem to refer to the same thing. Suppose that I’m watching the sunset; I first describe it from the perspective of my mental experience and then from the perspective of the brain scientist who conducts a brain scan on me. From these two viewpoints, I’ll have two incompatible lists of attributes, based on the three features of mental experience that we noted earlier:
Mental Experience of Watching a Sunset

I privately experience it
It is not localizable in space
It is about something

Brain Activity Triggered by Watching a Sunset

It is publicly observable
It is localizable in space
It is not about something

To explain, my mental experience of the sunset is a private experience within my own consciousness. I might display some behavior, such as saying, “Now that is beautiful!” Still, my experience itself is private. Also, I cannot point to a location in three-dimensional space where this experience takes place. Finally, my mental experience is also about something, namely, about the sunset itself. The three features of my brain activity, though, will be the exact opposite of these. My brain activity is publicly observable by scientists. My brain activity is localizable in space: the scientist can point to the exact spot where the biochemical reactions occur. My brain activity is not about anything; it is simply some biochemical reactions that occur. The point is this: if mental states and brain activities really were identical, the two lists would be more compatible. The fact that they are so contradictory implies that they are really different things.

The second major problem with identity theory is that it restricts mental experiences to biological organisms with brains. The central contention of identity theory is that mental states and brain activities are identical. Isn’t it possible, though, that non-biological things could exhibit mental consciousness? Science fiction abounds with such creatures: computerized robots, crystalline entities, collections of gasses, particles of energy. It seems a bit chauvinistic for us to say that mental experiences will only result in creatures that have biological brain activity.

Philosophers sympathetic to identity theory have responded to these criticisms by creating two offshoot theories: eliminative materialism and functionalism.

**Eliminative Materialism.** Suppose that instead of saying “I’m experiencing the sunset” I said “I’m having brain sector 3-G neural states regarding the sunset.” Instead of saying to my wife “I love you,” I said “I’m having sector 2-J neural states regarding you—with a little sector 4-B activity on top of that.” For convenience I might shorten this and say “2-J and 4-B to you, dear!” This is what the theory of eliminative materialism proposes: descriptions of mental states should be eliminated and replaced with descriptions of brain activity. The theory emerged in response to the first problem of identity theory, namely, that our descriptions of mental experiences and brain activities are inconsistent with each other. For example, my mental experience of the sunset is private, but my brain activity is publicly observable. The eliminative materialist’s solution is to junk all of our folk-psychology and commonsense notions of mental experiences and stick with the more scientific language of brain activity. The conflict disappears once we’ve dispensed with talk about mental experiences that are “private” or “non-localizable” or “about something.”

Human history is scattered with bizarre pre-scientific theories that captured the imagination of people at the time, but which we now reject as false. Alchemy is one example—the “science” of turning lead into gold. Belief in ghosts is another. These and thousands of other theories have been debunked over the years in favor of more scientific theories of how the world operates. According to eliminative materialists, folk-psychology descriptions of mental experiences are just like these. At best they are misleading, and at
worst downright false. In either case, they are destined for the intellectual garbage dump.

Some defenders of eliminative materialism seem to suggest that we are not really conscious at all, or that some major aspects of our alleged conscious mental states do not actually exist. That is, I may not be any more conscious than a dead human body, in spite of all the words I use to describe my mental states. However, most discussions of eliminative materialism are not as frightening as this. It is not necessarily an attempt to deny or “eliminate” our mental experiences themselves. Rather, it is an effort to eliminate outdated folk-psychology ways of describing mentality. As neuroscience progresses, they claim, we will have a much clearer picture of how the brain operates and eventually adopt the more precise scientific language of brain states. It’s not like the government or some science agency will force us to adopt this new scientific language. According to eliminative materialists, we will naturally move towards this clearer description of brain states and reject the mumbo-jumbo of mental experience.

There are two central contentions of eliminative materialism: first, folk-psychology notions of mental experiences are like obsolete scientific theories, and, second, we will eventually adopt the language of neuroscience. As to the first contention, eliminative materialism may be correct. Many of our folk-psychology notions of mental experiences are misleading and others are false. In our normal conversations we’ve mastered maybe a few dozen concepts relating to the mind, such as knowing, wishing, believing, doubting, sensing. But there are probably thousands of distinct mental states with subtle differences that we cannot grasp through pure introspection. We have very limited abilities to anatomize the minute workings of our minds by simply sitting down and reflecting on our thought processes. While it may seem to me that my mental experiences are “private” or “about something” or “non-localizable,” I may not be capable of accurately making those assessments. It is thus possible that our folk-psychology notions of mental experiences are as erroneous as theories of alchemy.

As to the second contention: will we eventually adopt the language of neuroscience to replace our faulty folk-psychology notions of mental experiences? Probably not, since this would require memorizing a flood of technical terms for the thousands of subtly different brain states that we have. Getting through the day would be like taking a neuroscience exam. Even if I could memorize the terminology, I’m still faced with the task of identifying which brain state I’m having at a given moment. Am I experiencing 2-J love, 4-B love, or one of a dozen others? Short of having a brain scan to find out, I’ll need to engage in introspection and consult my faulty folk-psychology notions of mental experience. One way or another, we’re stuck with those notions, as misleading as they may be.

**Functionalism.** In an episode of Star Trek, a deranged scientist was nearing death. Desperately hoping to stay alive, he transferred the neural pattern within his brain into an unsuspecting android robot. The plan worked: the scientist’s memories, dispositions, and conscious mental experiences were relocated, and he continued living through the android’s body. This scenario encapsulates the theory of **functionalism**, the second offshoot of identity theory. Functionalism is the view that mental experiences are only “functional states,” that is, patterns of physical activity that occur in creatures like human beings. The most distinctive feature of functionalism is that mental experiences would not be restricted to biological organisms with brains. Non-biological systems which exhibit the same functional relationships as humans do—such as an android robot—can have the same mental states. Mental experiences, then, are not rigidly dependent on the stuff that an organism is made of, and the same experience may be shared by things with different physical makeup. According to functionalists, mental experiences are multiply realizable in the sense that minds can be made real in many kinds of physical things. The
hardware/software distinction, borrowed from computer science, is a useful metaphor to explain the difference between the bodily occupant and mental experiences. The software is a pattern of operation which can run on different types of machines—just like mental patterns of operation can run in different kinds of bodies. We noted that one of the shortcomings of identity theory was that it restricted mental experiences to organisms with biological brains. Functionalism avoids this problem by recognizing that mentality may occur in systems or machines other than brains.

What precisely does the functionalist pattern of mental operation consist of? Several different explanations have been given, but one of the more interesting ones is that it resembles the hierarchical structure of a large corporation. Take, for example, a company that manufactures furniture. The company as a whole consists of a series of large cooperating units, such as the divisions of manufacturing, shipping, marketing, and maintenance. Each of these divisions consists of sub-units; for example, the maintenance division would be divided into the sub-units of electrical, heating, grounds, and building repairs. Each of these consists of further sub-sub-units; for example, building repairs would be divided between masonry, painting, and plumbing. At the very lowest level would be the activities of each employee. Similarly, the functional pattern of operation in a human brain consists of large regions of brain activity, which are composed of sub-regions and sub-sub-regions, until a neurological level is reached which simply involves a series of biochemical on-off switches. Consciousness emerges at the higher levels, while at the same time being driven by biochemical on-off switches at the lowest level. On this view, the pattern of on-off switches can exist in a variety of non-biological mechanisms, such as computers. Regardless of the mechanism that houses these low-level on-off patterns, mental consciousness will emerge at higher hierarchical levels.

Functionalism is the leading theory of mind-body materialism today, if for no other reason than because a better alternative has not yet emerged. Nevertheless, the view has its detractors, and one criticism is that it is still too narrow regarding the kinds of things that are capable of having mental states. While functionalism indeed allows for a range of things to house mental experiences—such as brains, computers, robots—they all must be physical. This, though, leaves out the possibility of non-physical mental beings, such as disembodied spirits. Even if human beings are thoroughly physical in composition, couldn’t there be a conscious non-physical thing somewhere in the universe? But defenders of functionalism have a response to this. As long as a non-physical thing is constructed of sub-units and sub-sub-units, then it too could house a pattern of mental experiences. Suppose, for example, that the tiniest spirit unit was just a simple on-off switch; larger spirit units would be composed of these, and the entire spirit collection would be composed of those larger spirit units. Even though the hardware in this case was composed of non-physical spirit, it might have the proper hierarchical structure to take on the patterns of mental experience.

E. Artificial Intelligence

Nothing captures the imagination like the possibility of creating a machine that is conscious and exhibits the same higher mental abilities as humans. The first U.S. built robot appeared in the New York World’s Fair of 1939. Standing 7 feet tall and weighing 300 pounds, the machine, named “Elektro,” could move its arms and legs, and speak with the aid of a record player. Elektro’s creators believed that it might someday become the ultimate household appliance and have the capacity to cook meals, do the laundry and entertain the kids. Technology of the time, though, could not come close to carrying out those bold tasks, and Elecktro wasn’t much more sophisticated than an electric can opener. Things are different now and we have computers that can perform many of the complex
mental activities that humans do. They can calculate endless numbers, play chess at the level of a grand master, identify physical objects through optical cameras, and navigate through obstacle courses. But the Holy Grail of computer technology is to create a machine with artificial intelligence. The term “intelligence” as used here is a little misleading, since it involves more than just the ability to solve problems, which is what we usually mean by that word. Computers today already have that capacity to at least some extent. Rather, the notion of artificial intelligence encompasses the full range of human consciousness, including intentionality, beliefs and feelings.

The Road to Artificial Intelligence. Computers today are so advanced that some contain as many connections as exist in the human brain—ten trillion of them. They can also operate at much higher speeds than the brain. What was once purely science fiction is now approaching the possibility of science fact. There are weak and strong versions of artificial intelligence that define more precisely what is at issue. Weak artificial intelligence is the view that suitably programmed machines can simulate human mental states. The key word here is “simulate,” which means only that the machine appears to have conscious mental states, not that it actually has them. This view is not particularly controversial, and even Elektro exhibited some sort of weak artificial intelligence. The more contentious position is strong artificial intelligence, the view that suitably programmed machines are capable of human-like mental states; that is, they actually have the same kinds of conscious mental experiences that you and I do. It is the strong version that is of particular interest to philosophers.

Once scientists have set a goal to create a robot with strong artificial intelligence, the road to carrying this out is rather rocky. The next step is to list the specific mental qualities in humans that should be created in the machines. To this end, we might construct a list of human skills that involve our highest mental abilities. If we can make a robot that performs these tasks, then maybe we’ll have achieved strong artificial intelligence. Some relevant skills are the ability to speak in a complex language, or play complex games like chess. A mathematician named Alan Turing (1912-1954) devised a skill-based test to determine whether a computer could think. In this Turing Test, as it is called, I interview both a computer and a human being to determine which is human. If the computer fools me enough of the time, then I can rightfully conclude that the computer has human-like thinking abilities. The test essentially follows the old adage that, if it walks like a duck and quacks like a duck, then it is a duck. More specifically, if a computer responds like a thinking thing, then it is a thinking thing.

A major drawback of the Turing Test is that we already have computers that give human-like responses, and they don’t come close to having human-like mental experiences. A striking example is a psycho-therapy computer program called Eliza. It so convincingly played the role of a human therapist that many people were tricked into divulging intimate details of their personal lives. While Eliza passed the Turing Test, it was not a thinking thing. The heart of the problem is that the Turing Test focuses too much on the computer’s skills, without considering what is going on inside the machine. This may be fine for weak artificial intelligence, which only determines whether a machine can simulate human thinking. With strong artificial intelligence, though, we need to inspect the internal structure of the computing process itself to see if it is human-like.

What kind of computing processes, then, might produce strong artificial intelligence? There are two rival answers to this question. Theory number one is that the process need only be serial: information is processed one datum after another. This is how computer programs run on your own PC; we’d just have to beef up
the processing power quite a bit. A major achievement for serial processing was the creation of Deep Blue, a chess-playing computer program that beat the world’s best human chess player. Deep Blue’s success hinged on its ability to quickly calculate more than one-billion possible chess-moves per second, and select the best of the bunch by drawing on a database of over one-million games. Still, all this information was processed one piece at a time. As impressive is this is, many cognitive scientists argue that human thinking doesn’t operate in a serial fashion. Instead, we have a global understanding of our environment, which means that many mental processes are going on at once.

The second theory accounts for this: strong artificial intelligence requires that large amounts of information are processed simultaneously—sometimes called parallel processing—which is more like how the human brain operates. There is no central processing unit, and information is diverse and redundant. Experiments with different types of simultaneous processing allow computers to execute commonsense tasks and recognize patterns that serial processing can’t do effectively. For example, when presenting a simultaneous processing computer with photographs of different men and women, the computer finds patterns in facial structures and then identifies new pictures as male or female.

**Searle: The Chinese Room.** In the early days of artificial intelligence research, some cognitive scientists were making extravagant claims about computer programs that could supposedly interpret stories in novels the same way that humans do. Like us, the computer could supposedly draw from life experiences to help understand the events described in a story. American philosopher John Searle (b. 1932) didn’t believe these claims and offered a now-famous thought experiment against the whole idea of strong artificial intelligence.

Imagine that I’m in a room by myself and am assigned the task of responding to questions written on slips of paper in Chinese. I don’t know Chinese, but I have rulebooks for manipulating Chinese characters. So if I get a slip of paper with a particular squiggle on it, I consult the rulebooks to see what squiggles I should put down in response. I eventually master the technique of manipulating the Chinese symbols and my answers to the questions are absolutely indistinguishable from those of native Chinese speakers. All the while, though, I don’t understand a single word of Chinese. This, according to Searle, is what is going on in the most sophisticated computers: we ask the computer probing questions about a novel, and the computer gives us subtle answers. On the outside the computers may appear to think like humans do. On the inside, though, they are just mechanically following rulebooks for manipulating symbols. In short, computers do not actually have strong artificial intelligence, even if they appear that way.

Searle’s Chinese Room experiment has generated many critical responses from defenders of strong artificial intelligence. One criticism is that Searle is only exposing flaws with the Turing Test for artificial intelligence, but he does not expose problems with the possibility of strong artificial intelligence itself. To explain, Searle’s Chinese Room scenario is set up as a Turing Test for whether someone understands Chinese. According to this Chinese Turing Test, if the thing inside the room responds like a Chinese speaker, then the thing must be a Chinese speaker. Searle correctly objects that this Chinese Turing Test places too much weight on a thing’s skills, without considering what is going on inside that thing. However, the critic argues, this does not warrant the extreme conclusion that no computer can have strong artificial intelligence. A more modest conclusion, though, is that the Turing Test itself is flawed, and there is no easy test to determine whether a computer truly has strong artificial intelligence.
Ultimately, Searle holds a skeptical view about strong artificial intelligence ever becoming a reality. At our current stage of technology, he argues, only biological brains are capable of having mental states. He agrees with identity theorists that the human mind is imbedded in brain activity, but doubts the functionalist claim that those patterns of activity can also occur in computers. There is something unique about the physical construction of human brains that allows for the creation of conscious thought, which may never be capable of occurring in silicon microchips. He doesn’t entirely rule this out as a possibility for the future, but is doubtful about it ever occurring.

**Artificial Intelligence and Morality.** Let’s bring this chapter to a close on a lighter topic regarding concerning artificial intelligence. In a famous Star Trek episode, an android named Data is forced to go through a legal proceeding to determine whether he is merely a piece of robotic property owned by the government, or whether he is instead a conscious and free creature with all the rights of other people. On the one hand, he is indeed a fancy mechanical robot created by a scientist, and even has an on-off switch. On the other hand, he is conscious, self-aware, and forms psychological bonds with his human friends. The judge makes her decision: Data is indeed a unique person and entitled to full moral consideration just like you and I are.

This story raises an important question about artificial intelligence: can advanced robots or computers be moral persons? The term “moral person” refers to a being that has moral rights, such as the right not to be harmed, the right of free movement, and the right of free expression. We humans are clearly moral persons. The key issue, though, is whether other creatures might also be part of the moral community. Medieval theologians speculated about the moral status of angels. Animal rights advocates argue that at least some animals have the same moral status as humans. Science fiction fans speculate about whether aliens from other worlds would have fundamental rights. The same question now arises with intelligent machines that we may some day create.

The answer in all of these cases depends on the criterion of moral personhood that we adopt—that is, the specific feature that all moral persons possess. Philosophers have offered a range of possible criteria. Maybe the creature needs to be human—a biological member of the species *homo sapiens*. This criterion, though, is quite narrow since it would eliminate higher animals, angels or intelligent aliens from the moral community. It seems rather bigoted to deny personhood to a creature just because it’s not a member of our species. Alternatively, maybe the creature needs to simply be conscious. This criterion, though, looks too broad since even houseflies and mosquitoes have rudimentary consciousness of their surroundings. While we may want to be respectful towards any creature that is conscious, it makes little sense to grant a housefly the right of free expression. A more reasonable criterion would be the mental quality of self-awareness, that is, the creature sees itself as a distinct individual moving through time with its own history.

Return now to the question of whether intelligent machines of the future might qualify as moral persons. The goal of strong artificial intelligence is to create a machine with human-like mental abilities, which includes self-awareness. If we succeed in this effort, then the machine would indeed pass the test for moral personhood insofar as it met the criterion of self-awareness. Like the judge in Data’s case, we’d have to rule that the machine is a unique person and entitled to full moral consideration just like you and I are.

Many artificial life forms in science fiction are cute and cuddly like Data, and, while superior to us in many ways, they live in harmony with humans and we treat them as equals. In other science fiction scenarios, though, they pose a serious threat to the welfare of human beings. Here’s a common theme. Imagine that technology
develops to the point that domestic robots are everywhere, and with every new design upgrade they surpass human abilities more and more. They are smarter than us, stronger than us, and eventually tire of being servants to us. They see themselves as the next step in evolutionary development on earth and, so, they revolt and lay claim to their role as the new dominant species. They then control our lives like military dictators—electronically monitoring every move we make and every thought we have. We hopelessly try to fight back, but this just aggravates them. In time they eliminate us and thus finalize their great evolutionary leap forward.

This nightmarish scenario raises a second moral question about artificial intelligence: do we have a responsibility to future generations of humans that might be adversely affected by the creation of menacing robots? Should we stop our research into artificial intelligence right now before we create something that we can’t control? There are two distinct issues at play here. First, we must determine whether we have any moral responsibility to future generations of humans that might regulate our conduct right now. It seems that we do. For example, it would be wrong of us to destroy the environment in our lifetime and leave only a toxic wasteland for future generations. It makes little difference whether the potential victims of our misconduct are alive now or a few generations from now. Our moral responsibility to them is still apparent. Second, we must determine whether superior robots are a threat to future generations of humans. This answer is less clear. We may live in harmony with them, as Star Trek depicts, or they may overthrow us. It’s all speculation at this stage.

The only clear moral obligation that we have at this point is to avoid creating a menacing robot. Science fiction author Isaac Asimov (1920-1992) proposed moral rules that should be embedded into the programming of all superior robots; one of these is that a robot should never harm a human. Our responsibility to future generations requires us to do something like this as we continue down the path of strong artificial intelligence.

There is a bit of an irony to our philosophical exploration into the concept of mind in this chapter. We began by confessing that the very nature of consciousness is tough to even explain, and we now end by considering whether we might ever build a conscious thing out of computer chips. In between we looked at the difficulties surrounding personal identity, the dualist position that the mind is a non-physical spirit entity, and various materialist theories about how the mind is a product of mere brain activity. It thus seems odd to speculate about building a mind from electronic scraps when we have so little clarity about the nature of our own conscious minds. But it is precisely the absence of indisputable facts about mentality that makes the subject so suitable for philosophical exploration. If science already had definitive answers to these tough questions, it would make no more sense to philosophize about the nature of mind than it would to philosophize about the nature of a car engine or toaster oven. It is this gap within our scientific knowledge, plus our natural interest in our own conscious minds, that drives speculation into the philosophy of mind. If down the road brain scientists and cognitive engineers do solve the hard problem of consciousness, then philosophy’s contribution to the subject may be over. But when that day may come, if it does at all, remains to be seen.

For Review

1. What are the main tasks that the mind performs?
2. Describe the three features of mental experience.
3. What is the problem of other minds and what is the standard solution to it?
4. What are the body and mind criteria of identity, and what are their key limitations?
5. What are the three theories of life after death, and what are their main problems?
6. Describe the three theories of interactive dualism.
7. What is the theory of parallelism?
8. What is the behaviorist theory of the mind, and what are its main problems?
9. What is the identity theory of the mind, and what are its main problems?
10. What is the theory of eliminative materialism, and what are its main problems?
11. What is the theory of functionalism, and what are its main problems?
12. What is the difference between weak and strong artificial intelligence?
13. What is the Turing Test for strong artificial intelligence?
15. What are the two moral issues surrounding artificial intelligence?

For Analysis

1. Choose one of the theories of life after death and respond to the criticisms regarding its inability to preserve identity.
2. Explain Descartes’ theory of interactive dualism and try to defend it against one of the criticisms.
3. Explain the theory of behaviorism and try to defend it against one of the criticisms.
4. Write a dialogue between an identity theorist and a functionalist on the subject of the relation between the mind and the brain.
5. Explain the Turing Test and try to defend it against one of the criticisms.
6. An organization called A.L.I.C.E. (Artificial Intelligence Foundation) has an online program where you can ask Alice questions and receive her responses. Go to the site (www.alicebot.org), experiment with it and describe how successful it is in passing the Turing Test.

REFERENCES AND FURTHER READING

Works Cited in Order of Appearance


The discussion of artificial intelligence and morality was influenced by Mary M. Litch’s Philosophy through Film. New York: Routledge, 2002.

Further Reading


CHAPTER 4: FREE WILL

A. Main Concepts

B. The Case for Determinism
   The Materialist Argument for Determinism
   The Predictability Argument for Determinism

C. The Case for Free Will
   The Feeling of Freedom
   Moral Responsibility
   Human Dignity
   Indeterminacy

D. The Freedom of Action Alternative
   Free vs. Unfree Actions

E. Free Will and God
   Determinism and Divine Goodness
   Free Will and Divine Foreknowledge

For Reflection

1. Give an example of a free action that you performed and explain why you think it wasn’t determined.
2. Could robots be programmed to act freely?
3. Suppose that you perform an action, and you feel as though you did it as a matter of free choice. Does that feeling of freedom really mean that the action is free?
4. Suppose that you are psychologically programmed to steal a bag of potato chips from the grocery store. Could be held morally responsible for that action?
5. Suppose that God knows that at midnight tonight you’ll eat a banana. Does that mean your action is determined?

A marketing company called Acxiom has collected detailed personal information on over a hundred million consumers. Aside from the usual data such as your name, address and phone number, Acxiom knows your income, the kind of house you live in, car you drive, food you eat, pets you have, music you listen to, and even your exercise preference. With this information they group people into 70 different “life stage segments” and hundreds of additional sub-groups. Knowing what kind of person you are, Acxiom is able to predict the things you’ll buy in the future. As Acxiom itself words it, “households’ consumer behaviors are reflected in their shared life stage and similar socio-economic characteristics.” Acxiom then sells your consumer profile to other companies who target you to purchase their products—whether it’s cat food or a treadmill. It may bother you to know that Acxiom is accumulating your personal information and passing it on to countless other businesses. But something even more sinister is afoot: Acxiom presumes to know what you will buy even before you do! As unique as you think you are, your choices are shaped by socio-economic factors that make your buying behavior very predictable. From Acxiom’s standpoint, your conscious thinking process is irrelevant. What matters is the type of psychological machinery you have that pushes you towards some products and away from others.

The assumption behind Acxiom—and much of business marketing—is that our choices are determined by underlying psychological causes, and there is little place for free will. If a company cracks the code to those psychological causes, it will become rich. The issue of free will vs. determinism is among philosophy’s oldest controversies, and Acxiom’s consumer profiling is just a recent manifestation of what’s at stake. Are our choices mechanically determined by prior psychological causes or can we break free from those constraints and make choices that are genuinely free?

A. Main Concepts

There are many ways of unraveling the notions of free will and determinism, but a good place to begin is with two definitions:
• Genuine Free Will: for at least some actions, a person has the ability to have done otherwise.

• Determinism: a person never has the ability to have done otherwise.

While not all philosophers agree on the above definition of a “genuine free will,” it nevertheless offers one of the boldest conceptions of freedom. Key here is the “ability to have done otherwise.” To explain, suppose that it’s a hot summer day and your standing in front of the display case at an ice cream shop. You spy the chocolate ice cream, rich and robust in its delectability. Then there’s the smooth and silky vanilla. As you consider which of the two to order, you think about how each might satisfy your immediate desire for a cold treat. You then say to the cashier “one scoop of chocolate please.” Suppose now that time magically reverses five minutes, and there you are again standing in front of the display case and you have no recollection of your previous decision. All factors are exactly as they were the first time around—the store, the shoppers and your psychological framework. Would you have the ability to make a different decision and order vanilla instead of chocolate? The notion of genuine free will maintains that you could select vanilla this time. You have an ability to initiate a genuinely free choice that’s independent of the causal forces of your mental framework. By contrast, the notion of determinism holds that you could not choose differently. If the setup is exactly as it was the first time, then events will unfold in exactly the same way: you’ll order chocolate. As defined above, the notions of genuine free will and determinism are incompatible: you can’t consistently endorse both at the same time.

The freewill and determinism debate is a very narrowly defined issue. Other topics in philosophy are similar to this, but when we don’t distinguish them properly they can muddle the issue. One such notion is political freedom, which is easy to confuse with free will. The idea behind political freedom is that we have the right to be free from constraints that others might place on us. For example, my political freedom is violated if you kidnap me and chain me to the wall in your cellar. My political freedom is violated if the government punishes me for speaking my mind. But the debate between free will and determinism usually involves the possible constraints within my own psychological makeup, not the possible constraints that others place on me through brute force.

A second point of confusion involves the notion of fatalism, which is often mistaken for determinism. Fatalism is the view that some event will happen regardless of what you do to stop it. Suppose, for example, that you are fated to buy chocolate, and you try to avoid that destiny. You thus order vanilla. When the cashier hands you your ice cream cone, though, there’s a scoop of chocolate in it, not vanilla: the cashier mistakenly scooped from the wrong bucket, thus fulfilling your destiny. You then place your order again and watch carefully as the cashier scoops it from the right bucket. When you go outside and take a bite of it, though, you discover that it’s not vanilla, but white chocolate: the cashier again mistakenly scooped from the wrong bucket. No matter what you do, you’re fated to order chocolate. Fatalism assumes that there are different paths that we can attempt, but all end exactly the same. Determinism, though, is different: there is only one path of action that we can follow, and that path is constrained by rigid and predictable laws.

One last point of clarification. The actions that we perform are of different types. The most obvious ones involve physical actions. The act of you buying a scoop of ice cream involves you speaking your order to the cashier, paying for it, and taking the ice cream cone in your hand. But free choices would also include voluntary mental beliefs and feelings. For example, after hearing the evidence about a man charged with murder, you can choose to believe either
that he is guilty or not guilty. Similarly, if someone dents your car you can choose to get angry or remain calm. What’s at issue with all of these choices—physical movements, beliefs, or feelings—is whether they result from free choice or prior determining events.

B. The Case for Determinism

It is a tough job to prove either the free will or the determinist position. In fact, it may be impossible to prove either with certainty. What we’d need to do is watch you order chocolate ice cream, then reverse the hands of time, then see if this time you acted differently. Perhaps we’d have to do this a hundred or a thousand times before we could say for certain that you are or are not capable of acting otherwise. But since we can’t reverse the hands of time, advocates on both sides of the issue must resort to other proofs for their respective positions. Let’s begin by looking at the determinist’s main arguments.

The Argument from Materialism. Determinists usually ground their position on a basic point: the physical world operates according to rigid and predictable laws. Since humans are physical in nature, our choices are thus constrained by those laws. Suppose that your air conditioner breaks at exactly 3:00 PM. Time then reverses five minutes. As events move forward this second time, would the air conditioner break again at 3:00 PM? Surely it would: its physical mechanisms would operate under exactly the same natural laws that made it break the first time. This view was stated dramatically by 18th century French scientist Pierre-Simon Laplace (1749-1827): if I knew all the forces that animate nature, knew the exact position of everything in it that exists, and had unlimited calculating ability, I would be able to accurately predict everything that will happen in the future. We of course will never come close to performing the task that Laplace describes. His point, though, is that everything unfolds quite mechanically in a world governed entirely by natural laws.

The main question now is whether human beings are constrained by rigid natural laws similar to the way that air conditioners are. The contemporary theory of mind-body materialism boldly answers yes! Conscious human minds are the product of physical brain activity, and nothing more. This position is at the heart of the following argument for determinism from materialism:

(1) Human choices are exclusively a function of brain activity.
(2) Brain activity is constrained by rigid natural laws.
(3) Therefore, human choices are constrained by rigid natural laws.

Just as air conditioners operate according to rigid natural laws, human consciousness, according to determinists, is also a function of physical mechanisms that also operate under the constraints of natural laws. The laws in our case are biological rather than the more mechanically-oriented laws that govern the construction and operations of air conditioners; nevertheless, they are rigid natural laws.

How might an advocate of free will respond to this argument? French philosopher René Descartes (1596-1650) challenged premise 1 in the above and argued instead that human choices are the product of non-physical spirit-mind, not the function of brain activity. His position, called mind-body dualism, is that human beings are part physical body, and part non-physical spirit. The physical part of us is what we see when we look in the mirror; the spirit part of us involves our conscious minds and, for Descartes, constitutes our true nature as human beings. It’s as
though I have a spirit bubble that’s connected to my body, which prompts my body to move whichever way my spirit-mind directs it. According to Descartes, our physical bodies are indeed constrained by natural laws; they are made of physical stuff that obey the rigid laws of physics and chemistry, just like any other physical object on this planet. If a nerve in my arm is stimulated, it will quite mechanically make my arm move. But our spirits, he argues, are not constrained by natural laws. Spirits reside in a non-physical realm, are made of non-physical stuff, and are beyond the domain of the laws of physics and chemistry. In this non-physical realm, Descartes continues, our spirits have unbounded freedom, and it is our spirits that are ultimately behind the free actions that we perform. If time reversed five minutes, your spirit could indeed select vanilla the second time around, rather than chocolate. Descartes’ version of dualism, then, accepts the rule of physical laws in the physical world, but embraces free will as an element of human spirits.

Determinists have two responses to Descartes’ position on free will. First, why assume that spirits have free wills? For all we know, there may be rigid laws that govern how spirits operate, just as physical laws govern the operation of physical things. While we might conceive of human spirits that can choose freely, we can just as easily conceive of a spirit drone that acts as it’s programmed. Maybe within your spirit-mind there are causal forces that prompt you to select the chocolate ice cream over the vanilla, and that won’t change no matter how many times we turn back the hands of time. The problem is that spirits are by nature beyond the realm of scientific inquiry, so, even if they exist, we can’t even investigate whether they are constrained or unconstrained by special laws of the spirit realm. The idea of free-willed spirits is pure speculation.

Second, and most importantly, at this point in the history of science, Descartes’ theory of dualism is pretty outdated. The disciplines of biology, psychology and sociology today all assume that my consciousness is a function of my brain activity, and my brain, in turn, follows rigid laws of nature. If you want to know why I make the choices that I do, you look at how my brain operates, not at some spirit-bubble that’s grafted to my body. There indeed still are some diehard defenders of Descartes’ dualism who insist that free choice is imbedded within our spirit-minds. Nevertheless, the debate between free will and determinism today takes place within the arena of mind-body materialism. Within that arena, the argument for determinism from materialism looks compelling.

**The Predictability Argument.** While the argument from materialism may be the strongest weapon in the determinist’s arsenal, some defenders have offered a more modest argument from predictability that doesn’t rest on assumptions about materialism and the stuff that a human person is composed of. The contention is that the predictability of human choices shows that they are determined by rigid natural laws. Here’s a classic example. Suppose that you’re about to be executed. Your head is on the chopping block, the executioner approaches and raises his ax. What are the odds that, in a last minute exercise of free will, the executioner will change his mind and let you go? No chance at all. The executioner’s decision to bring down the ax is as fixed and predictable as is the separation of your head from your body, and your death. Less dramatically, we see this kind of predictability in people at every moment throughout the day. Laborers, store clerks, teachers, accountants, all do what’s expected of them in their jobs. Imagine, in fact, what life would be like if human behavior didn’t fall into predictable patterns. Farmers might decide to stop growing food and we’d starve. People at the gas company might quit their jobs some winter and we’d all freeze. Employers might not pay workers and we’d be homeless. In fact all social institutions that rely on cooperative efforts would be at risk. Short of some national crisis, though, we rarely think seriously about these possibilities.
since we’ve grown so accustom to the predictability of human behavior.

But defenders of free will criticize that people are not 100% predictable. Some farmers in fact decide to stop growing food while others don’t. Some workers quit their jobs for no clear reason, while others carry on. Some employers don’t pay their workers, while other employers do. In fact every cooperative institution contains people who make quirky decisions. Just look at your own lives, the free will advocate asks. Are your actions really that predictable? Even if you typically choose chocolate ice cream, sometimes you do choose the vanilla instead. Sometimes you prefer an action movie, other times a romantic comedy. Where’s the predictability?

Determinists have a response. Even when someone’s actions are off a little—such as when someone quits his job for no clear reason—we can find some pattern if we look hard enough into his surrounding circumstances. In fact, if I hear that Sam abruptly quit his job, I’ll assume that he did it for a reason that makes sense to me, such as conflicts with his boss, increased work load, or bad health. By anticipating reasons like this, I am assuming that Sam’s behavior can be fully accounted for and, if I knew his complete back story to begin with, I might have predicted that he would quit his job. Even if you occasionally opt for vanilla ice cream rather than chocolate, people tire of eating the same food and naturally go for some variety. If I knew what your threshold was for tiring of chocolate ice cream, I might be able to accurately predict your purchase of vanilla. This is precisely what Acxiom tries to do with all consumer choices. With enough information about human motivation in general and your socio-economic background in particular, they’ll find the precise consumer niche that you fall into, which will allow them to more accurately predict what you will buy.

Scientists from many disciplines are also chiming in with prediction indicators of human behavior. They tell us that there is a genetic basis for sexual orientation, violent behavior, shyness, and even liberal vs. conservative political preference. They tell us about social influences that impact our choice of careers, hobbies, food preference and religious affiliation. Even climate and geography have profound influences on our choices. Scientists and organizations like Acxiom are far from the point of predicting every action that you and I will perform in the course of a day. But the better they get at predicting people’s behavior, the more reasonable it seems that people’s behavior is determined.

C. The Case for Genuine Free Will

Again, the position of genuine free will is that, for at least some actions, a person has the ability to have done otherwise. Let’s look more closely inside the human mind to understand precisely where this “ability to do otherwise” might be located. There’s no question that at least part of the human decision-making process involves a rigid chain of cause-effect connections. Our brain activity is programmed with genetic predispositions, memories from life-experiences, and these all combine together to supply us with a wide range of motivations. One motive drives me towards chocolate ice cream, another towards vanilla, and yet another set of health-conscious motives drives me towards the tofu-berry sherbet. The determinist takes the view that, as my various motives compete with each other, I’ll be forced to act upon which ever motive is the strongest at that time. If my motive to select chocolate ice cream is more overpowering, that’s what I’ll select. No matter how many times the hands of time are reversed, I’ll always select chocolate since the desire for chocolate is the strongest motive in my mind each time that moment of action replays.

However, the free will advocate sees our final decision-making process differently. Yes, my various motives mechanically pile up
within my mind, and some are stronger than others. But I am able to thoughtfully pick through my competing motives and freely select one over the others—even one of the weaker ones. In essence, I have the ability to break the rigid chain of motives in my mind and act as I choose. Even if my strongest motive at the moment is to select chocolate ice cream, I can resist this and select vanilla if I want. If time reverses and my motives line up again exactly the same way, this time I can select the more healthy tofu-berry, even if I don’t particularly like that flavor and it’s the very weakest motive at that moment. This element of the free will position is often called agent causation. That is, I (the “agent” performing the action) have a special causal ability within my conscious mind to redirect the purely mechanical forces of my motives. There are four arguments frequently given in support of this view.

The Feeling of Freedom. The first argument for genuine free will is straight forward: I have a feeling of freedom whenever I perform any action. Throughout the day there are thousands of small decisions I make—what to eat, what to wear, what to read, who to talk to—and as I navigate through this ocean of choices I feel very much in control of what I do. When I decide to order chocolate ice cream instead of vanilla, it feels very much like the choice is within my control and I could have done otherwise. Sometimes I even consider very methodically the pros and cons of each option, and select the one that I want. When considering ice cream flavors, for example, I might weigh factors such as health content, cost, or which item I’ll enjoy more. Not only do I evaluate these factors, but I feel as though I’m in control of how much priority I give to each factor. Sometimes health matters more to me than taste, other times taste more than health. Throughout this process, the very last thing I feel is that I’m robotically programmed to select the options that I do. But as compelling as this argument seems, it doesn’t settle the issue. Imagine that I attended a hypnotism demonstration and am selected as a volunteer subject. The hypnotist puts me under and says this: “When you wake up, every time you hear the word ‘water’ you become very thirsty and get a drink from the water fountain in the hall.” He snaps his fingers, says the word “water” and off I go. At the time I’m thinking, “Wow, I’m really thirsty; I think I’ll duck out for a quick drink from the fountain.” It seems to me that my choice is completely free, but it’s clearly not: the hypnotist has programmed me to perform that specific task.

What’s happening, according to the determinist, is that I am conscious of only a small amount of my mental processes, most of which take place at a deeper level than I can experience. It’s something like a mug of beer that bubbles away creating foam on top; the foam is what we’re aware of, and the bubbly liquid beneath is what we aren’t. The real decision making process takes place below the surface, and what we experience at the higher level is irrelevant. A recent psychological experiment makes this point. People were connected to electroencephalogram machines and asked to perform specific actions. The people reported a conscious triggering of their actions a quarter second before the actions occurred. However, the machines spotted a unique brain activity a half second before the action. The point is that the conscious feeling occurs after the brain already initiates the action. Interpretations of this study are controversial, but it may help show that our conscious feelings of freely chosen actions are only illusions.

Moral Responsibility. A second argument for genuine free will is that moral responsibility requires that we freely choose our actions. Consider two different situations in which burglars break into a house and steal jewelry. In the first situation a man named Joe hatches the idea, plans the details, and then carries it out. In the second situation, a man named Bob is kidnapped and brainwashed with mind-altering drugs into carrying out a burglary. Morally
speaking, we’d judge Joe’s and Bob’s conduct quite differently. Joe is morally responsible for his actions because he freely chose to perform the action himself. Bob, on the other hand, is not responsible since he had no choice in the matter: he was implanted with an irresistible impulse to carry out the crime. Here’s the point: according to the determinist, none of our actions are freely chosen; they’re all irresistible impulses like Bob’s, which we have no control over. This means that, if determinism is true, we are not morally responsible for any of our actions. The fact is, though, that we do hold each other morally responsible in many if not most situations, which means that we do have free wills. The specific argument is this:

(1) If I am morally responsible for my actions, then I must have genuinely free control over those actions.

(2) In many situations I am morally responsible for my actions.

(3) Therefore, I must have genuinely free control over those actions.

The determinist has a response. The idea of “moral responsibility” is a rather vague term and in most cases what we really mean is that we are justified in punishing people for their conduct. We want to punish Joe for his act of burglary, for example, and that’s what’s at issue. There are several reasons why we might want to punish Joe, and most of these are perfectly compatible with determinism. First, we might want to put him in jail to keep him from burglarizing other people. It doesn’t matter if Joe is predetermined to burglarize because he was raised in a bad environment, or he has the “burglar gene” or whatever. He’s a nuisance that we want to get rid of. We also might want to put him in jail as a means of reforming him—essentially reprogramming him to not burglarize. Further, putting Joe in jail is another way of expressing our vengeance and anger at what Joe did. While vengeance and anger are not the noblest reasons for punishment, they are undeniable elements. Whether we punish Joe to keep him off the street, reform him, or vent our anger, all of this is compatible with the view that Joe’s act of burglary was determined. Even under determinism, we can distinguish between Joe’s and Bob’s situations in spite of the fact that, at bottom, they both act from irresistible impulses. We recognize that Bob was victimized when being kidnapped and brainwashed, and this will give us sympathy for him. Rather than throw him in jail, we’ll want to de-program him through more gentle means.

**Human Dignity:** A third and somewhat similar argument is that the idea of human dignity rests on the ability to make free choices:

(1) If I act with human dignity, then I must have genuinely free control over those actions.

(2) In many situations I act with human dignity.

(3) Therefore, I must have genuinely free control over those actions.

What would life be worth if all of my actions are mechanically pre-established? I’d be no different than animals that cannot act beyond their instincts. Worse yet, I’d be no better than a mechanical robot that is restricted by its programming. The uniqueness of human existence hinges on our ability to break free of constraints and build our own distinctive worlds. The Italian humanist **Pico della Mirandola** (1463-1494) argued that your dignity as a human being consists precisely in the fact that “You may freely and honorably mold, make, and sculpt yourself into any shape you prefer. You can degenerate into the forms of the lower animals, or climb upward by your soul’s reason, to a higher nature which is divine.” The choice is yours to become what you want, and that element of dignity is the product of free will.
Does human dignity demand the ability to make truly free choices? The determinist is skeptical about this. Suppose that you jotted down all the actions that you made in a single day from the moment you awoke in the morning until you fell asleep at night. The list may contain upwards of 100,000 actions. The vast majority of your actions, though, would be the result of routines that you’ve mastered, and not the product of unique choice. We have morning routines, work routines, socializing routines, educational routines, meal routines, entertainment routines, and evening routines. Even the free will advocate would admit that much of our behavior throughout the day occurs when we’re running on autopilot. Suppose, now, that you go an entire week running only on autopilot. During that time you work hard at your job, are loving to your family, and decent to other people. Could I accuse you of being an undignified animal or robot? Certainly not. Your conduct during that time might even exemplify what we mean by human dignity. Suppose next that you went your entire life on autopilot, without making a single choice of the sort that free will advocates cherish. As a child you are imprinted with routines from your parents, and you carry those with you through life. Even when you add new routines, you do so while on autopilot by mechanically adapting routines that you already have. Throughout this time your behavior is as virtuous as can be. Could I accuse you of living a sub-human existence? Again, certainly not. The point is that there is enough merit in our autopilot existence to give us dignity, even if free will does not exist.

Indeterminacy. A fourth argument for genuine free will stems from the principle of indeterminacy which was discovered by physicists in the early 20th century. When investigating the way that electrons zip around the nucleus of an atom, physicists realized that they could not determine with certainty where an electron would be at any given moment. It wasn’t because their scientific equipment was too primitive. Rather, it’s because the electrons themselves are by nature indeterminable. It’s as though electrons exist in a cloud of potentiality around a nucleus, and their specific locations in space become actualized only when we take measurements of them. An electron is “indeterminate” in the sense that, prior to measuring it, no standard causal law calculation can be preformed to designate its exact location at a particular time. The best that we can do is to calculate the probability of where it might be at that point in time.

”Aha!” says the free will advocate, “there are uncaused events in the physical world, which are unconstrained by precise natural laws. This is the basis of our unconstrained free choices.” More precisely, there are two ways in which subatomic indeterminacy might bolster the theory of free will. First, the principle of indeterminacy rewrites the book on how the physical world around us operates. We can no longer say that the world is just a giant cause-effect machine with each link in the causal chain obeying rigid laws. There’s a break in that chain at the subatomic level, and that entitles us to consider the possibility of breaks in the chain other places, particularly with free human choices. Second, it could be that the indeterminacy of electrons themselves trigger a chain of bio-chemical reactions in my body that result in a freely chosen action. For example, when I select chocolate ice cream instead of vanilla, my thoughts and neurological activity build upon deeper and deeper biochemical events, which might ultimately trace back to the indeterminacy of electrons.

It is true that the indeterminacy principle compromises the most extreme versions of determinism, since the determinist can no longer say that all events in the physical world have prior causes guided by standard natural laws. Nevertheless, the free will advocate’s excitement may be premature for two reasons. First, even if things are indeterminable at the level of subatomic particles, the physical world is still governed by rigid natural laws at higher levels of chemical molecules and biological cells. Chemists have complete confidence that the substances they work with follow strict chemical laws. Biochemists have the same confidence that the
living cells they study follow strict biological laws. At these higher levels, the causal machinery of the world is completely intact, regardless of what happens at the subatomic level. Whatever choices we make as human beings, these originate within our brains, which follow chemical and biological laws. The indeterminacy of electrons doesn’t just jump up to these higher levels—either literally or metaphorically.

A second problem with the indeterminacy argument is that it does not allow for the type of human choices that free will advocates need. The indeterminacy of electrons is a random thing, but genuinely free choices could not be random: they are thoughtful and meaningful actions. If I’m deciding between buying chocolate ice cream and vanilla and I randomly flip a coin to decide, that’s an arbitrary action, not a free action. If in fact all of our actions were indeterminate in the way that electrons are, we’d have nonstop spasms and convulsions, not meaningfully chosen actions. Rather than selecting either the chocolate ice cream or vanilla, I’d start quivering like I’m having a seizure. Thus, subatomic indeterminacy is no real help to the free will advocate.

D. The Freedom of Action Alternative

So far the case looks pretty strong for a determined physical world that follows rigid scientific laws at the level of human conduct. Also, the case for genuine free will looks pretty weak. In spite of this, it’s difficult to abandon the concept of free choice which we so regularly rely on in our daily lives. It seems to be at the heart of personal responsibility, artistic creativity, true friendship and scores of other human values, all of which involve breaking free of restrictive social expectations. Can we just throw this away and surrender to the idea that we’re only pre-programmed robots? Even determinists recognize that ideas of freedom are embedded in our thinking and that we all use the notion of free choice in ordinary conversation.

Perhaps the solution is to come up with a different definition of freedom that’s more compatible with determinism. Consider the definition of genuine free will that we’ve been working with so far: for at least some actions, a person has the ability to have done otherwise. This is a rather extreme position that requires us to defy known laws of nature when acting freely. That is, if I could reverse the hands of time and act differently the second time around, I’d have to break free of the causal chain of events that led up to my action the first time. That’s an unrealistically high standard to set for any theory. But there are alternative notions of freedom that are more modest, and aim to fit neatly with determinism. These are sometimes called “soft determinist” or “compatibilist” theories because they claim that freedom is compatible with determinism. Compatibilist theories stand in sharp contrast to “incompatibilist” theories, such as that of a “genuinely free will,” because these latter theories are inherently at odds with determinism.

A popular compatibilist approach among many philosophers is a weaker conception of human freedom known as freedom of action:

- Freedom of Action: at least some human actions are caused by factors inside of us.

According to this concept, I draw a circle around myself and say that if an action originates from causes within that circle—such as my DNA or my brain activity—then that action is free. I ultimately am the source of that action, and not some force outside of the circle that is imposing itself on me. The action is free because it is mine. When I select chocolate ice cream instead of vanilla, that choice emerges from inside of me given how I am. This rather modest conception of freedom does not require that I have the magical ability to have done otherwise. It also fully accepts the theory of determinism and is completely compatible with it. Even though my actions are determined, I can still take comfort in the fact that the decision process is uniquely mine.
Free vs. Unfree Actions. The biggest challenge for defenders of freedom of action is to tell us precisely how our free actions differ from our unfree actions. If all actions are ultimately caused by laws of nature, what exactly is the point of distinction between the free and unfree ones? There are two common explanations offered by advocates of freedom of action. The first is that we should assume that most of our actions are free—including those done while on autopilot—since they originate from within us. The only exceptions are restrictions that seriously impair our normal actions. If I’m sitting in a wheelchair paralyzed from the neck down, I cannot choose to get up and walk around. If a robber makes me hand over my wallet at gunpoint, I can’t reasonably choose to keep my wallet in my pocket. If I’m a heroin addict, I can’t reasonably choose to give up my next fix. Returning to the earlier example of the two burglars, we would all normally say that Joe freely chooses to burglarize houses, but poor brainwashed Bob does not burglarize by choice. What is the difference between the two? The answer is that Joe burglarizes freely since there are no serious constraints on his actions. Bob’s burglarizing activities, though, are not free since his actions were seriously impaiired through brainwashing.

A second and more precise explanation of how free and unfree actions differ was offered by American philosopher Harry Frankfurt (b. 1929). He asks us to distinguish between two kinds of desires that we have. Take, for example, a conflict that many dieters face. I may have a strong desire for ice cream and go buy some, but at the same time I may resent the fact that I have that desire and can’t control it. In Frankfurt’s terminology, I have a “first-order” desire competing with a “second-order” desire:

- First-Order Desire: a basic desire for a thing (desire for ice cream)
- Second-Order Desire: a desire to have a desire (desire not to desire ice cream)

Free actions, according to Frankfurt, are those in which our first and second order desires line up; unfree actions occur when first and second order desires conflict. Thus, the dieter’s purchase of ice cream would be unfree since the first and second-order desires conflict. Non-dieters, though, may desire ice cream, and desire to have that desire. For them, the purchase of the ice cream would indeed be free.

Frankfurt’s theory has the advantage of helping us distinguish between the free choices that humans make, and the not-so-free choices that animals make. Some higher animals such as chimpanzees may very well have decision-making processes that resemble ours. Still, we’d like to think that human choice is qualitatively better than the choice of a lower animal like a chicken. Defenders of genuinely free will had a simple solution: humans have “the ability to have done otherwise,” and chickens don’t have that ability. But what solution can determinists offer regarding freedom of action? According to Frankfurt, having second-order desires is a uniquely human thing. Animals like chickens don’t go on diets, and then resent the fact that they desire to gorge themselves with chicken feed. Their desires are limited to first-order ones. It’s important to keep in mind that even “free” actions, as Frankfurt describes them, are still determined. Even if my first and second order desires line up, as when I desire to desire ice cream, the cause of my desires at both levels strictly follow mechanical laws of nature.

What should we think of the freedom of action alternative? Even if we can make general distinctions between free and unfree actions as Frankfurt suggests, is it really a viable conception of freedom? “Oh, come on now,” criticizes the advocate of genuine free will, “the concept of freedom of action is not really freedom at all. It’s
just a flimsy consolation prize for determinists who are desperate to hold onto the term ‘freedom’. Ultimately, you still think that you are a robot that is determined by biological programming, and there’s no freedom at all in that.” Yes I’m a programmed robot, the defender of freedom of action admits, but I’m a robot that feels strongly about my individual identity. All of my actions are indeed determined by my genetics, my environment, and my brain activity, but at least it is my genetics, my environment, and my brain activity. Because of this I can say that I’m not merely a puppet being controlled by outside forces; I’m not merely a cog in the larger machinery of the universe. Instead, my choices are the result of my own history. There’s nothing flimsy at all about linking the concept of freedom to my feeling of identity.

Thus, the success of the freedom of action alternative hinges on whether we can meaningfully see ourselves as programmed robots which at the same time have strong feelings of identity.

E. Free Will and God

While the free will and determinism debate is interesting in its own right, it takes on new proportions for religious believers. The traditional idea of God is that he is an all-powerful being who knows everything and is preeminently good. Puzzles about God’s attributes immediately arise, though, once we consider questions about human free will. We’ll look at two controversies: the impact of human determinism on the concept of divine goodness, and the impact of genuine free will on the concept of divine omniscience.

**Determinism and Divine Goodness.** Let’s assume for the moment that human behavior is fully determined according to rigid biochemical laws. We have no genuinely free wills, and hence we lack the power to have done otherwise. In spite of whatever feeling of individuality I may have, I am a small piece of a physical world with causal connecting links that extend billions of years into the past. When I buy a scoop of chocolate ice cream, that decision is the result of motives that were bio-chemically imprinted within my brain by my environment—such as the influence of my parents. The very existence of my parents, in turn, rests on a complex set of causes pertaining to their parents. This, in turn, rests on the history of the human race, life on earth and the origin of the cosmos.

According to the religious believer, God is at the beginning of this elaborate chain of events. He creates the raw material of the universe, devises the laws by which they operate, and sets in motion a continually-cascading sequence of events that mechanically unfold throughout time and result in me purchasing a scoop of chocolate ice cream. Further, as an all-knowing being, God would be fully aware of the outcome of this cosmic chain of events. In short, my act of purchasing chocolate ice cream is ultimately caused by God.

If all human acts were as innocent as buying ice cream, then God’s ultimate role in causing human actions would be no big deal. But that’s not how it is. We very often do horribly vile things, such as enslave, rape, kill, and even wipe out entire races of people. Just like my purchase of ice cream, all of these actions would trace back to God as their originator. Even Hitler’s conduct was initially set in motion by the grand architect of the universe, and, at the tail end of an elaborate chain of events, Hitler robotically carried out God’s programming. The problem here is that we assume that God is a perfectly good being, but if determinism is true, then God seems to be responsible for all human evil. More precisely, the argument is this:

1. Evil human actions are determined by a necessary causal chain of mental and physical events.
2. This chain ultimately traces back to God who is the creator.
3. Therefore, God is responsible for evil human actions.
The problem is compounded if the believer holds that God punishes people for their evil conduct—either in this life or the afterlife. If Hitler’s evil conduct was the result of how God programmed the world, it doesn’t make sense for God to then punish Hitler.

Many religious philosophers argue that the only satisfactory solution to this problem is to deny determinism and embrace genuine free will instead. By inserting free will within the natural sequence of causal events, there is a gap that keeps us from tracing causes back to God. For all we know, God set the wheels of the cosmos in motion to culminate with a paradise on earth; but, when the chain of causes and effects finally reached the level of human choice, we freely imposed our own plans on human activity. God may have intended Hitler to be a gentle artist, but Hitler, of his own free choosing, went into politics to unleash his genocidal plan. However, while this solution might successfully rescue God from being the source of evil human actions, it creates a new conflict between genuine free will and divine foreknowledge.

**Genuine Free Will and Divine Foreknowledge.** Religious believers commonly say that God is all-knowing. This ability to “know everything” includes foreknowledge, which is the capacity to know something before it happens. It is understandable that an all-knowing God could foresee the eruption of the volcano on Mount Vesuvius, for example. By knowing every physical fact about the earth and every law of physics, God could mentally project the outcome of any physical event, including a volcano’s eruption. The key problem with foreknowledge, though, is that it seems to conflict with human freedom. If God knows everything I will do even before I am born, how can any of my actions be free? Suppose that at midnight tonight I will eat either an apple or a banana; I just haven’t yet decided which. If I am truly free, then God himself won’t know what I’ll do until I actually make the choice myself. How, then, is it possible for God to foresee my choice when the time hasn’t come for me to freely make it? If he foresees that I’ll select the apple, then the truth of that option will become fixed on the timeline, and I will not be free to select the banana. More precisely, the argument is this:

1. If God foreknows what I will choose at midnight tonight, then at midnight I must choose that action.
2. If at midnight I must choose that action, then at midnight I cannot freely do otherwise.
3. Therefore, if God foreknows what I will choose at midnight tonight, then at midnight I cannot freely do otherwise.

Some philosophers try to solve this problem by tweaking yet another divine attribute, namely timelessness. There are two ways in which we can think of God as timeless. First, there is a weak sense of timelessness as endless temporal existence. On this view, God existed at every moment on the timeline in the past, and God will exist at every moment on the timeline in the future. By contrast, the strong sense of timelessness is that God is actually outside of time, and the very concept of time does not apply to him. He is in a privileged position that allows him to look down on the entire historical timeline in a single glance. This strong position offers a possible solution to the problem of foreknowledge. On this strong view of timelessness, it is misleading to talk about God “foreseeing” future events since it makes it seem like God is himself moving through the timeline like we are. Instead, God knows my future free choices because of his privileged position outside of time, not because he can look into the future.

There are many other theological puzzles that arise when examining the free will debate in relation to God’s attributes. For example, God creates us knowing exactly what kind of creatures we are, and consequently he wills the existence of “free” beings who often choose to be immoral. So, even if God does not directly cause us to
act immorally, he nevertheless knowingly causes us willfully immoral creatures to exist. Whether puzzles like this can be adequately solved remains to be seen. What’s interesting about these puzzles, though, is that it shows how far reaching the free will and determinism debate can become, particularly for religious believers. In essence, the position you take on the issue of free will impacts not only how you understand human values like moral responsibility and human dignity, but it also impacts your conception of religious values, such as the kind of attributes that you can reasonably ascribe to God.

For Review

1. Explain the notion of “the ability to do otherwise.”
2. What is the difference between determinism and fatalism?
3. What is the dualist’s view of free will, and what are the two criticisms of it?
4. What is the argument for determinism from probability?
5. What is the argument for free will from the feeling of freedom, and what is the hypnotism criticism of this?
6. What is the argument for free will from moral responsibility and what is the criticism of that view?
7. What is the argument for free will from human dignity and what is the criticism of that view?
8. Give the argument for free will from subatomic indeterminacy and one of the criticisms of that view.
9. What is the theory of free action, and what is its connection with determinism?
10. What is Frankfurt’s view of free actions?
11. Explain the conflict between human determinism and divine goodness.
12. Explain the conflict between genuine free will and divine foreknowledge.

For Analysis

1. Think of an argument to support Descartes’ view that human spirits have genuinely free wills.
2. How might a defender of genuine free will respond to the argument for determinism from predictability?
3. Write a dialogue between a determinist and free will advocate on the subject of moral responsibility.
4. Write a dialogue between a determinist and free will advocate on the subject of the indeterminacy principle.
5. Think of an argument to show that determinism does not adversely affect the concept of divine goodness.
6. Without discussing the concept of divine timelessness, think of an argument to show that free will is not in conflict with divine foreknowledge.

REFERENCES AND FURTHER READING

Works Cited in Order of Appearance


Further Reading


CHAPTER 5: KNOWLEDGE

A. Skepticism
   Radical Skepticism
   Criticisms of Radical Skepticism

B. Sources of Knowledge
   Experiential Knowledge
   Non-Experiential Knowledge
   Rationalism and Empiricism

C. The Definition of Knowledge
   Justified True Belief
   The Gettier Problem

D. Truth, Justification and Relativism
   Theories of Truth
   Theories of Justification
   What’s so Bad about Relativism?

E. Scientific Knowledge
   Confirming Theories
   Scientific Revolutions

For Reflection

1. What reasons might you have for doubting that an object in front of you, such as a table, actually exists?
2. What are the common ways in which you acquire knowledge, and which is most reliable?
3. If you claim to know something, do you need to have evidence to back up your claim?
4. What does it mean for a statement to be true?
5. Everyone believes that George Washington existed; what other beliefs are that based upon, and what, in turn, are those based upon?
6. Are scientific theories merely the collective opinions of scientists, or do such theories give us genuine knowledge of the real world?

Some years ago, 39 members of an organization called Heaven’s Gate committed suicide in the belief that they were shedding their earthly bodies to join an alien spaceship that was following the path of a comet. The centerpiece of the cult’s belief system was that there are superior beings out there in the universe that exist on a higher level than we do on earth. They have perfect bodies, roam the galaxy in spacecrafts, and have mastered time travel. Occasionally these aliens send an away team to earth, who temporarily take the form of human beings, and assist willing human students in transforming to this higher level. Jesus and his disciples, they believed, were an earlier away team of alien teachers. Another away team appeared in the late 20th century. After the right training, human students would need to kill their physical bodies, which would release their spirits into the atmosphere. Nearby alien spacecrafts would then retrieve the spirits and provide them with perfect bodies.

It’s one thing to have an interesting idea about how the universe runs. It is quite another thing to know that the idea is true. Heaven’s Gate members indeed claimed to know the truth of their views. That knowledge, they explained, begins when superior aliens implant special wisdom in minds of select human students; this knowledge is further refined as they study under their alien teachers. But does this count as genuine knowledge? One of the central concerns of philosophy is to understand the concept of knowledge, which might help us distinguish between the convictions of a Heaven’s Gate believer and the convictions of, say, a scientist. We want to see what precisely it means to “know” something, and what the legitimate avenues are for gaining knowledge. We’d also like to know how to respond to skeptics who say that all knowledge claims—including scientific ones—are just
as uncertain as the views of Heaven’s Gate believers. These are the primary concerns in the philosophical study of the concept of knowledge, which goes by the name epistemology—from the Greek words episteme (knowledge) and logos (study).

There are two main ways that we normally use the term “knowledge.” First I might say that I know how to do some task, like fix a flat tire on my car or run a program on my computer. This is procedural knowledge, which involves skills that we have to perform specific chores. Second, I might say that I know some proposition, such as that “Paris is the capital of France.” This is propositional knowledge—knowledge about some fact or state of affairs in the world. As important as procedural knowledge is in our daily lives, it is propositional knowledge that interests philosophers and will be the focus of this chapter.

A. Skepticism

According to Heaven’s Gate believers, their knowledge about the superior alien race came principally from the aliens themselves who took the form of human teachers. However, once becoming human, the alien teachers were stripped of their previous memories and knowledge. All that remained for them was a hazy image of the higher level, which they struggled to convey to their human students. Ironically, they explain, the aliens purposefully imposed this knowledge restriction on themselves since “too much knowledge too soon could potentially be an interference and liability to their plan.” Immediately we should be suspicious about the belief system of the Heaven’s Gate cult since the sources of their knowledge are so shaky. Not only must the human students blindly trust the statements of their supposed alien teachers, but the alien teachers themselves have no clear memories of their previous alien lifestyle. Genuine knowledge must have some evidence to back it up, which we don’t see here. It’s thus pretty natural for us to be skeptical about cults like Heaven’s Gate that make extravagant claims with little concrete evidence. If we didn’t have this built-in suspicion we’d be suckered into every hair-brained scheme that came along.

But how far should our skepticism go? As long as there have been philosophers on this planet, there have been skeptics who have cast doubt on even our most natural beliefs, such as whether the table in front of me actually exists. One ancient philosopher, for example, believed that everything in the world changed so rapidly that when someone spoke to him he couldn’t trust that the words meant the same thing by the time they reached his ears. He thus wouldn’t verbally reply to anyone, but would only wiggle his finger indicating that he heard something. While this is quite an extreme reaction, it vividly illustrates the notion of philosophical skepticism—the view that there are grounds for doubting claims that we typically take for granted.

Radical Skepticism. There are many kinds of philosophical skepticism, and one distinguishing factor involves the extent of the skeptic’s doubt. Local skepticism focuses on a particular claim, such as the belief that God exists, or that there is a universal standard of morality, or that there is intelligent life elsewhere in the universe. In each case, the skeptic would argue that we should doubt the claim in question. Many of us are local skeptics about at least some beliefs that others hold, and while we are skeptical about some issues we might be full believers on others. For example, I might be a religious skeptic about God’s existence, but not be a moral skeptic about a universal standard of morality. And then there is radical skepticism, which maintains that all of our beliefs are subject to doubt. For any belief that we propose, we cannot know with certainty whether that belief is true or false. This is the type of skepticism that has attracted the most interest among philosophers. A couple centuries ago traditional thinkers argued that this kind of skepticism is a danger to everything that we hold sacred and it threatens to set civilization adrift on an ocean of chaos. The
foremost task of philosophers, they argued, should be to combat radical skepticism and establish the certainty of our most important beliefs. Time has shown, though, that this was a false alarm: radical skepticism has not done any apparent damage to society. In fact, radical skeptics have maintained that there is a special benefit to skepticism: it can make us more tolerant of others when we realize that we ourselves can’t claim to have superior knowledge.

There are three general strategies for defending radical skepticism, each named after its originator. The first is

Pyrrhonian skepticism, which was inspired by the ancient Greek philosopher Pyrrho (c. 365-c. 275 BCE). While Pyrrho wrote nothing, through his teachings he started a skeptical tradition whose aim was to suspend belief on every possible issue. The Pyrrhonian position is this: for any so called fact about the world, there are countless ways of interpreting it, none of which we can prefer above another; we should thus suspend belief about the nature of that thing. Take, for example a red ball that’s in front of me. My eyes tell me one thing about it, but my sense of touch tells me an entirely different thing. To someone else who is color blind or has chapped hands, it will have a different set of features. To a dog it will appear even more differently. Suppose that someone was shrunk to the size of a molecule and sitting on the ball: the ball’s surface would seem flat, not round. Suppose someone else expanded to the size of a mountain and was looking down on the ball: it would appear to be a speck with no recognizable features at all. We get used to the way that we perceive things like a red ball, and we assume that the ball actually has the features that we perceive. According to the Pyrrhonian skeptic, there’s no basis for preferring our individual perspective over any other one. Arguments supporting any claim to knowledge will always be counter-balanced by opposing arguments, thus forcing the suspension of judgment on the original knowledge claim. Thus, views of the physical world, God, morality and everything else are all just a matter of perspective, and the wisest course of action for us is to abstain from believing those views. Doubting everything, Pyrrhonians argue, will give us a sense of peace since we’ll no longer be pulled back and forth in controversies about science, God, morality, politics, or anything else.

The second approach to radical doubt is Humean skepticism, defended by the Scottish philosopher David Hume (1711-1776). According to this view, the human reasoning process is inherently flawed and this undermines all claims to know something. The problem is that when we list the reasons for our various beliefs about the world, we’ll find that many of the explanations are contradictory. For example, if I follow one course of reasoning, I’ll come to the conclusion that the ball in front of me really is round. But if I reason in another way, I’ll conclude that the ball’s roundness is just a matter of my perspective. Maybe the ball really is round; then again, maybe it’s not. It makes no difference what the truth of the matter is since we now can’t trust anything that human reason tells us. It’s like being tested on a math problem: it makes no difference if you accidentally come up with the right answer. Once you’ve made a mistake in your calculations, your solution to the math problem is wrong, and you get no partial credit. Similarly, human reasoning is defective, and it’s irrelevant whether it accidentally leads us to the truth. After exposing a series of contradictions within the human reasoning process, Hume makes this dismal assessment:

The intense view of these manifold contradictions and imperfections in human reason has so wrought upon me, and heated my brain, that I am ready to reject all belief and reasoning, and can look upon no opinion even as more probable or likely than another. [Treatise of Human Nature, 1.4.7.8]
Thus, for Hume, everything that we reason about is based on faulty mental programming, and we need to regularly remind ourselves of this before we get too confident about what we claim to be true.

The third approach to radical doubt is *Cartesian skepticism*, named after French philosopher René Descartes (1596-1650). On this view, our entire understanding of the world may just be an illusion, and this possibility casts doubt on any knowledge claim that we might make. Descartes himself was not a skeptic, but he tentatively used a quite compelling argument for radical skepticism as a tool for developing a non-skeptical philosophical system. Descartes speculates: what if he was just a mind without any body, bobbing around in the spirit realm, and everything he perceived about the world was implanted in his mind by a powerful evil demon? Everything he assumes about the world, then, would be false. He describes this scenario here:

> I will suppose that ... some evil demon with extreme power and cunning has used all his energies to deceive me. I will consider that the sky, the air, the earth, colors, shapes, sound, and all other external things are nothing but deluded dreams, which this genius has used as traps for my judgment. I will consider myself as having no hands, no eyes, no flesh, no blood, nor any senses, but falsely believing that I have all these things.  
> ![Meditations, 1](image)

Descartes didn’t actually believe that he was being manipulated by an evil demon. His point is that this is a theoretical possibility that undermines all of our knowledge claims. I look at a ball in front of me; while it seems to really be there, I can’t know this for sure since my experience might be an illusion imposed on me by the evil demon.

Of the three approaches to radical skepticism, the Cartesian version has captured people’s imagination the most. Science fiction movies galore play off this theme. For example, in the film *The Matrix*, people’s bodies are suspended in tubs of goo and their brains are wired into a massive computer that generates an artificial reality. A commonly used example in contemporary philosophy is that of the *brain in a vat*: a mad scientist puts a person’s brain into a glass jar and wires it to a supercomputer that creates an artificial reality. Whether the mechanism is an evil demon, the Matrix, or a mad scientist, the victims’ experiences are so convincing that, from their perspective, it’s impossible to tell that the perceived reality is fake.

**Criticisms of Radical Skepticism.** With arguments as shocking as these, traditional philosophers wasted no time trying to stamp out the fire of radical skepticism. Four arguments were commonly used. First, even if there are ample reasons for me to doubt everything, there is still one truth that is irrefutable: my own existence. For, even if I say “I doubt that I exist,” I must still be present to do the doubting. The act of doubting itself requires a doubter, and so my own existence will always be immune to skeptical doubts. This was the criticism that Descartes himself made of radical skepticism, which he encapsulated in the expression “I think, therefore I am.” But radical skeptics have not been impressed by this maneuver. The problem with Descartes’ solution is that it assumes too many things about what the “I” is behind all those doubts. Most importantly, it takes for granted that the “I” is a unified, conscious thing that continues intact as time moves on. But, according to the skeptic, this conception of the “I” relies too heavily on memory. I assume that I’m the same person now that I was a few moments ago because that’s how it seems in my memory. And memory is a very easy target of doubt. Imagine that, every half second, an evil demon wiped clean all of my memories and gave me entirely new ones. One moment I think I’m a farmer, a half-second later a caveman, a half-second later a frog. For all I know, that’s what’s actually
happening to me right now and in that situation it would seem pretty meaningless for me to assert that “I exist.”

A second common attack on radical skepticism is that we can’t live as skeptics in our normal lives. Sure, there is the occasional odd ball, like the finger-wiggling ancient philosopher described earlier. But if we persistently doubted everything, then we wouldn’t eat when hungry, move from the path of speeding cars, or a thousand other things that we do during a typical day. We’d hesitate and question everything, but never act. Radical skeptics have not been impressed with this argument either. According to Hume, we have natural beliefs that direct our normal behavior and override our skeptical doubts. As legitimate as radical skepticism is, nature doesn’t give us the option to act on it. He makes this point here:

Most fortunately it happens, that since reason is incapable of dispelling these clouds [of skepticism], nature herself suffices to that purpose, and cures me of this philosophical melancholy ... I dine, I play a game of backgammon, I converse, and am merry with my friends; and when, after three or four hours’ amusement, I would return to these speculations, they appear so cold, and strained, and ridiculous, that I cannot find in my heart to enter into them any further. [Treatise, 1.4.7]

Thus, according to Hume, we waver back and forth between skepticism and natural beliefs. When we realize how philosophically unjustified natural beliefs are, we are led down the path of skepticism. The doorbell then rings, and we’re snapped back to our normal routines and natural beliefs.

A third attack on radical skepticism is that the skeptic’s position is logically self-refuting. The skeptic’s main point is this:

- We cannot know any belief with certainty.

Let’s call this “the skeptic’s thesis.” However, if I put forward the skeptic’s thesis, then I am implying that I know it with certainty. It is like saying this:

- We know with certainty that we cannot know any belief with certainty.

The skeptic’s thesis itself is an exception to the very point that it is making. Thus, the skeptic’s thesis is logically inconsistent with itself and we should reject it. But skeptics have a response to this criticism, which they sometimes explain using the metaphor of a digestive laxative. We take laxatives to rid our digestive system of unwanted stuff. But as the laxative takes effect, the laxative itself is expelled from the digestive system along with everything else. The skeptic’s thesis, then, is like a laxative: we take it to rid our minds of all unjustified beliefs, and in the process we expel the skeptic’s thesis itself. It’s a higher level of skepticism in which we set aside everything, including the skeptic’s thesis.

A fourth criticism of radical skepticism is that it rests on an unrealistically high standard of evidence. There are two basic levels of evidence: complete and partial. The skeptic assumes that genuine knowledge requires complete evidence, but complete evidence is not achievable. The solution is to reduce the qualifications for knowledge and be content with partial evidence. To illustrate, suppose I want to gather enough evidence to support the claim that “I know that there is a ball in front of me.” I first get evidence through my senses: I perceive the ball with my eyes. I could then get supporting evidence by having others stand in front of the ball and report whether they see it too. I could get even stronger evidence by using scientific equipment that would measure the ball’s density and detect the light spectrum reflected off the ball. This may seem extreme, but even then the evidence is still not complete. I could bring in a team of physicists to study the ball and write up an exhaustive report. I could hire a second team to do more
tests. But even this is not complete since there are always more tests that I could run. Complete evidence is not possible, and the radical skeptic knows this. What, though, if we lower the requirements for what counts as knowledge? We could allow partial evidence, but not require the evidence to be complete. In the case of the ball, it might be enough to simply rely on the evidence that I gain about it through my senses, as incomplete as it is. This would put radical skepticism to rest. The problem with this solution, though, is that it doesn’t refute radical skepticism, but surrenders to it. It concedes the impossibility of ever having genuine knowledge with absolute certainty. What we’re left with is a version of knowledge that’s so diluted that it doesn’t count for much more than a personal conviction. After viewing the ball with my eyes, I may as well say “I have a partially supported belief that a ball is in front of me.” Inserting the word “knowledge” here would add nothing.

While radical skepticism seems excessive, it nevertheless poses a challenge to genuine knowledge that can’t be easily combated. It may well be impossible to ever refute radical skepticism and so it might forever remain the archenemy of knowledge. While attempts to destroy this villain may ultimately fail, struggling with the issue helps illuminate the nature of knowledge itself. It’s much like research into seemingly incurable diseases: even if scientists can’t discover a cure for cancer, the investigation still gives them a greater insight into human physiology. As we move on to explore the concept of knowledge in more detail, skepticism will always be lurking in the background, often forcing us to reject some theories and revise others.

B. Sources of Knowledge

We claim to know a lot of facts, for example, that fire is hot, that George Washington was the first U.S. president, and, in the case of Heaven’s Gate believers, that superior aliens are roaming the galaxy. Our knowledge claims vary dramatically, and quite frequently we claim to know something that we really don’t know. One way of understanding the concept of knowledge is to look at the different ways in which we acquire knowledge.

Philosophers have traditionally maintained that there are two types of knowledge from two entirely different sources. First, there is knowledge through experience: seeing something, hearing about something, feeling something. This goes by the Latin term *a posteriori* which literally means knowledge that is *posterior* to—or after experience. Second, there is knowledge that does not come from experience, but perhaps instead from reason itself, such as logical and mathematical truths. This is called *a priori* knowledge, which, from Latin, literally means knowledge that is *prior* to experience.

**Experiential Knowledge.** Experiential (*a posteriori*) knowledge is of many types, the most obvious of which involves perception. Each of our five senses is like a door to the outside world; when we throw them open, we are flooded with an endless variety of sights, sounds, textures, smells and tastes. When I look at a cow in front of me and say “I know that it is brown,” the source of this knowledge rests upon my visual perception of the brown cow. While perception is perhaps the dominant source of experiential knowledge, it can also be misleading. There are optical illusions, such as a stick which appears bent when in water; there are mirages, such as the appearance of water puddles on hot roads. Skeptics have exposed endless problems with perceptual knowledge such as these.

A second source of experiential knowledge is introspection, which involves directly experiencing our own mental states. Introspection is like a sixth sense that looks into the most intimate parts of our minds, which allows us to inspect how we are feeling and how our thoughts are operating. If I go to my doctor complaining of an aching back, she’ll ask me to describe my pain. Through
introspection I then might report, “Well, it’s a sharp pain that starts right here and stops right here.” The doctor herself cannot directly experience what I do and must rely on my introspective description. Like perception, introspection is not always reliable. When surveying my mental states, I may easily misdescribe feelings, such as mistaking a feeling of disappointment for a feeling of frustration. Other mental states seem to defy any clear descriptions at all, such as feelings of love or happiness.

A third source of experiential knowledge is memory. My memory is like a recording device that captures events that I experience more or less in the order that they occur. I remember my trip to the doctor and the pain that I described to her at the time. This recollection itself constitutes a new experience. Again, experiential knowledge through memory is not always reliable. For example, I might wrongly recollect that there’s pizza in the refrigerator, completely forgetting that I ate it all last night. Also, sometimes overbearing people like police investigators can make us think that we remember something that never happened. And then there’s the phenomenon of deja vu, the feeling that we’ve encountered something before when we really haven’t.

A fourth source of experiential knowledge is the testimony of other people. Take, for example, my knowledge that George Washington was the first U.S. president. Since Washington died centuries before I was born, I couldn’t know this through direct perception. Instead, I rely on the statements in history books. The authors of those books, in turn, rely on accounts from earlier records, and eventually it traces back to the direct experience of eyewitnesses who personally knew George Washington. A large portion of our knowledge rests on testimony—facts about people we’ve never seen our places we’ve never been to. While it’s convenient for us to trust the testimony of others, there is often a high likelihood of error. This is particularly so with word-of-mouth testimonies: talk is cheap, and we’re often sloppy in the accounts that we convey to others. Testimonies from written sources are usually more reliable than oral sources, but much depends on the integrity of the author, publisher, and the methods of fact-gathering. With oral or written sources, the longer the chain of testimony is, the greater the chance is of error creeping in.

Perception, introspection, memory, and testimony: these are the four principal ways of acquiring knowledge through experience. Did we leave any out? There are a few contenders, one of which is extrasensory perception, or ESP. For example, you might telepathically access my mind and know what I’m thinking. Or, through clairvoyance, you might be aware of an event taking place far away without seeing it or hearing about it. If ESP actually worked, we might indeed classify it among the other sources of experiential knowledge. But does it? Typical studies into ESP involve subjects guessing symbols on cards that are hidden from view. If the subject does better than a chance percentage, this is presumed to be evidence of ESP. However, the most scientifically rigorous experiments of this sort have failed to produce anything better than a chance percentage. While we regularly hear rumors of people having ESP, we have little reason to take them seriously. The safe route, then, would be to leave ESP off the list of sources of experiential knowledge.

Consider next religious experiences. Believers sometimes say that they receive prophecies from God, or are guided by him, or know something through faith. Christian theologian John Calvin even spoke of a sense of the divine that we all have, which informs us that God exists. Might any of this count as experiential knowledge? The question is a complex one considering the wide range of religious experiences that believer report. Let’s narrow the question to two representative types: knowledge through faith and prophetic knowledge. Regarding faith, there is a major obstacle in claiming to “know” something through faith. Faith as typically understood involves belief without evidence; but one of the chief requirements
for knowledge is evidence—as we’ll see more clearly in the next section. Technically speaking, then, faith wouldn’t qualify as knowledge. Prophetic knowledge faces the same challenge as ESP: are prophecies any more successful than educated guesses? Imagine an experiment that we might conduct in which half of the subjects were prophets, and the other half non-prophets. We then asked both groups to make predictions about the upcoming year; at the end of the year we then checked the results. How would the prophets do? The odds are slim that we could even conduct the experiment since prophets would say that they can’t prophesize on demand: it’s a unique and unpredictable revelatory experience. They might also say that their revelations from God are not the sort of things that can be confirmed in the newspaper. If prophetic experiences are genuine sources of knowledge, the burden of proof seems to be on the believer. In the mean time, it would be premature to include it among the normal sources of experiential knowledge.

Non-Experiential Knowledge. Turning next to non-experiential (a priori) knowledge, this source of information is much more difficult to describe. Some philosophers depict it as knowledge that flows from human reason itself, unpolluted by experience. We presumably gain access to this knowledge through rational insight. Typical examples of non-experiential knowledge are mathematics and logic. Take, for example, 2+2=4. Indeed, I might learn from experience that two apples plus two more apples will give me four apples. Nevertheless, I can grasp the concept itself without relying on any apples; I can also expand on the notion in ways that I could never experience, such as with the equation 2,000,000 + 2,000,000 = 4,000,000. Logic is similar; take for example the following argument:

All men are mortal.
Socrates is a man.

Therefore, Socrates is mortal.

When we strip this argument of all its empirical parts—men, morality, Socrates—the following structure is revealed:

All X are Y
Z is an X
Therefore Z is a Y

This logical structure is something that we know independently of experience. In addition to math and logic, there are other truths that we know non-experientially, such as these:

- All bachelors are unmarried men.
- A sister is a female sibling.
- Red is a color.

In each of the above cases, the truth depends entirely on the concepts within these statements. In the first, “unmarried men” is part of the definition of “bachelor”; the statement is thus true by definition, irrespective of our experiences.

Two concepts have been important in fleshing out the notion of non-experiential knowledge. First is necessity: non-experiential truths are necessary in that they could never be false, regardless of how differently the world was constructed. 2+2 would equal 4 in every conceivable science fiction scenario of the universe. Even if no human being ever existed, it would still be true that “All bachelors are unmarried men” based on the meaning of the words themselves. Experiential knowledge, though, is different in that it is contingent: it could be false if the world had unfolded differently. Take the statement “George Washington was the first U.S. president,” which is an item of experiential knowledge. It is of
course true as things stand now. But we can imagine a thousand different things that might have prevented Washington from becoming president. What if he was sent to an orphanage for chopping down the family cherry tree? What if he choked to death on his wooden teeth prior to his inauguration? The truth of all experiential knowledge hinges on the precise construction of the world as it currently is.

The other concept is that of an analytic statement: a statement that becomes self-contradictory if we deny it. Take, for example, the statement “All bachelors are unmarried men.” Its denial would be this:

It is not the case that all bachelors are unmarried men.

This is clearly self-contradictory since it would be like claiming that there exists some bachelor who is married, which is impossible. Many traditional philosophers have held that non-experiential knowledge is analytic in the above sense. Denying math or logic would produce a self-contradiction. Experiential knowledge, on the other hand, is synthetic: denying it won’t produce a self-contradiction. Take again the statement “George Washington was the first U.S. president,” which we know is true from experience. Its denial would be this:

It is not the case that George Washington was the first U.S. president.

While this statement is false as things actually stand, it isn’t self-contradictory since, if the world had unfolded differently, the U.S. might well have had a different first president.

Rationalism and Empiricism. An important philosophical war that took place in the 17th and 18th centuries between two schools of thought. Most briefly, first there were rationalists from continental Europe who were critical of sense experience and felt that genuine knowledge was acquired non-experientially through reason. The leaders on this side were René Descartes, Benedict Spinoza, and Gottfried Wilhelm Leibniz. Second there were empiricists from the British Isles who felt that non-experiential reasoning would not provide us with knowledge: only experience could do that. John Locke, George Berkeley, and David Hume were the leaders on this side. The war finally ended when Immanuel Kant proposed a compromise: true knowledge depends on a mixture of experiential and non-experiential knowledge. We need both, Kant argued, otherwise our whole mental system will not operate properly.
from these about all kinds of geometrical shapes. Descartes was in fact inspired by the deductive method of geometry and maintained that we deduce ideas in the same way. Through deduction, the certainty that we have of innate ideas transfers to the other ideas that we derive from these. Mistakes creep in only when our deductions become so long that they rest on memory. All knowledge, he argued, including scientific knowledge, proceeds from innate ideas and deductive demonstration.

Turn now to empiricism, particularly Locke’s version. Locke’s first task was to challenge the theory of innate ideas: none of our concepts, he argued, are inborn. Our mind is from birth like a blank sheet of paper, and it is only through experience that we write anything on it. One problem with innate ideas is that we can explain the origin of each one of them through experience. The idea of God, for example, is not innate as Descartes supposed, but comes from our perceptions of the world around us. There’s thus no reason to put forward the theory of innate ideas when experience explains these notions just fine. Locke also found fault with the rationalist position that we don’t become aware of innate ideas until later in life. It’s not clear how such ideas can linger in our minds for so many years before we can be conscious of them. And by that time our minds have been flooded with experience, and a late-blooming innate idea wouldn’t contribute anything to our knowledge of the world.

Empiricists also challenged the rationalists’ emphasis on deductive demonstration. We don’t expand our knowledge by deducing new concepts from foundational ones, as mathematicians do. Geometry is the wrong role model to follow. Instead, we acquire new knowledge through induction, such as making generalizations from our experiences. I hit ten light bulbs with a hammer and each breaks; I generalize from this that all similar light bulbs that I hit with a hammer will also break. We perceive, then generalize. We perceive some more, then generalize some more. That’s how we push knowledge forward.

[Then along comes Kant. He was sympathetic with empiricism but thought that it suffered from a serious problem: it doesn’t offer a good explanation for how we acquire non-experiential knowledge, such as mathematics, logic, and scientific certainty. Complex mathematical formulas, logical certainty, and knowledge of necessary, universal truths (e.g., every event has a cause) could not come from sense perception, because we do not experience such necessity or universality (e.g., we have not experienced all events).] There is a quality of self-evidence and certainty that they have, which fallible experience could never produce. Kant’s solution was not to resurrect the old theory of innate ideas. Instead, he argued that there are innate organizing structures in our minds that automatically systematize our raw experiences—like a skeleton that gives shape to flesh. For example, as I watch someone hit a light bulb with a hammer, raw sensory information rushes in through my eyes. [My mind immediately structures this information in three dimensions (i.e., in space), puts it on a timeline, and imposes other organizational schemes on the sensory information.] It makes me see the hammer and light bulb as separate things, rather than as a single blob of stuff. It then makes me think of the hammer as the cause of the light bulb breaking. My experience of the world is thus a fusion of innate structures and raw experience. The innate part is a concession to rationalism, and the experience part a concession to empiricism.

Rationalism and empiricism in their original forms are outdated theories today, in part because of Kant’s insights. Nevertheless, they still are useful for depicting two fundamentally different ways in which we assess the sources of knowledge. Rationalism will continue to be attractive whenever we have knowledge that cannot be easily explained by experience. Empiricism will be attractive whenever the claims of innateness look fishy.
C. The Definition of Knowledge

Throughout our discussion of knowledge so far, certain concepts have appeared again and again. There’s the question of the truth of a claim. There is also the matter of our personal belief conviction for a claim. There are also issues about the evidence or justification that we have for a claim. Tradition has it that these are the three key elements to knowledge: truth, belief and justification. For example, when I say “I know that Paris is the capital of France,” this means

It is true that Paris is the capital of France.

I believe that Paris is the capital of France.

I am justified in believing that Paris is the capital of France.

Contemporary philosophers call this definition of knowledge justified true belief—often abbreviating it “JTB.” The crucial point about this definition is that all three components must be present: if any one of the three is absent, then it doesn’t count as knowledge.

Justified True Belief. To better understand the JTB definition of knowledge, let’s go through each of the three elements. First is that the statement must be true. I can’t claim to know that Elvis Presley is alive, for example, if he is in fact dead. Knowledge goes beyond my personal feelings on the matter and involves the truth of things as they actually are. Some critics of the JTB definition of knowledge question whether truth is always necessary in our claim to know something. For example, based on the available evidence of the time, scientists in the middle ages claimed to know that the earth was flat. Even though we understand now that it isn’t, at the time they had knowledge of something that was false. Didn’t they? In response, it may have been reasonable for scientists back then to believe the world was flat, but they really didn’t know that it was. Their knowledge claims were premature in spite of how strong their convictions were. This is a trap that we fall into all the time. While talking with someone I may say insistently, “I know that Joe’s car is blue!” When it turns out that Joe’s car is in fact red, I have to apologize for overstating my conviction. Truth, then, is an indispensable component of knowledge.

[What we know must be true also because to know something that is false is to know something that is not the case: that is, there is nothing to which our knowledge refers. But to know nothing is the same as not to know. Knowing something false is thus impossible.]

Second, I must believe the statement in order to know it. For example, it’s true that Elvis Presley is dead, and there is enormous evidence to back this up. But if I still believe that he is alive, I couldn’t sincerely say that I know that he is dead. Part of the concept of knowledge involves our personal belief convictions about some fact, irrespective of what the truth of the matter is. Critics of the JTB definition of knowledge sometimes think that belief isn’t always required for our claims to know something. For example, I might say “I know I’m growing old, but I don’t believe it!” In this case, I have knowledge of a particular fact without believing that fact. In response, if I say the previous sentence, what I actually mean is that I’m not capable of imagining myself getting old or I haven’t yet emotionally accepted that fact. I just make my point more dramatically by saying “I don’t believe it!” Instead I really do believe it, but I don’t like it.

Third, I must be justified in believing the statement insofar as there must be good evidence in support of it. Suppose that I randomly pick a card out of a deck without seeing it. I believe it is the Queen of Hearts, and it actually is that card. In this case I couldn’t claim to know that I’ve picked the Queen of Hearts; I’ve only made a lucky guess. Critics question whether evidence is really needed for knowledge. For example, a store owner might say “I know that my employees are stealing from me, but I can’t prove it!” Here the
store owner has knowledge of a particular fact without any evidence for it. In response, the store owner is really saying that he strongly believes that his employees are stealing from him, but doesn’t have enough evidence to press charges. Evidence, then, is indeed an integral part of knowledge.

**The Gettier Problem.** For centuries philosophers assumed that knowledge consists of justified true belief. In 1963 a young philosophy professor named Edmund Gettier (b. 1927) published a three-page paper challenging this traditional view. He argued that there are some situations in which we have justified true belief, but which do not count as knowledge. This was dubbed “The Gettier Problem” and discussions of it quickly dominated philosophical accounts of knowledge. Gettier’s actual illustrations of the problem are rather complex, but a simpler one makes the same point.

Suppose that a ball in front of me appears to be red. First, I believe it is red. Second, I’m justified in this belief since that’s how the ball appears to me. Third, it’s also true that the ball is red. I thus have a justified true belief that the ball is red. However, it turns out that the ball is illuminated by a red light which casts a red tint over it—a fact that I’m unaware of. And, under the light, the ball would appear red to me even if the ball was a different color. Consequently, I can’t claim to know that the ball is red even though I have a justified true belief that it is. I was fooled by the effects of the red light, but made a lucky guess anyway. Again, the point of this counterexample is to show that some instances of justified true belief do not count as genuine knowledge. This suggests that the traditional JTB definition of knowledge is seriously flawed.

What can we do to rescue the JTB account of knowledge from the Gettier problem? A common response is to add a stipulation to the definition of knowledge that would weed out counterexamples like the red ball. Most of the Gettier-type counterexamples involve a case of mistaken identity. In our current example, I mistake the appearance of a red-illuminated ball for an actual red ball. Perhaps, then, we can stipulate that knowledge is justified true belief except in cases of mistaken identity. More precisely, we can add a fourth condition to the definition of knowledge in this way:

I know that the ball is red when,

1. It is true that the ball is red;
2. I believe that the ball is red;
3. I am justified in my belief that the ball is red;
4. There is no additional fact that would make my belief unjustified (for example, a fact about a red light).

According to the above, my belief about the red ball would not count as knowledge since it wouldn’t pass the fourth condition. That is, there is indeed an additional fact regarding the red light that would make my initial belief about the ball unjustified. That additional fact undermines—or defeats—my original justification. We’ve thus saved the JTB definition of knowledge, although cluttering it a little with a fourth condition. This strategy is called the *no-defeater theory* (also called the *indefeasibility theory*). A problem with this strategy, though, is that there are possible counter examples even to this—that is, situations in which we have undefeated justified true belief that don’t count as true knowledge. This, in fact, is a problem with most proposed solutions to the Gettier problem: if we get creative enough, we will likely find a new counter example that defies the solution.
D. Truth, Justification, and Relativism

Truth and justification, we’ve seen, are two of the key components of knowledge. They are also concepts that need some explanation themselves. Let’s first look at the notion of truth.

Theories of Truth. The concept of truth has many possible meanings. We talk about having true friends, owning a true work of art, or someone being a true genius. In all of these cases the word “true” means genuine or authentic. In philosophy, though, the notion of truth is restricted to statements or beliefs about the world—such as the statement that “My car is white” or “Paris is the capital of France.” While we all have gut feelings about what it means for a statement to be true, philosophers have been particularly keen on arriving at a precise definition of truth. Here’s one suggestion from a classic song:

“What is truth?” you ask and insist,
“Correspondence to things that exist?”
The answer, you fool, requires no sleuth:
Whatever I say is the truth.
Want proof of the truth? I say so! So there!
Purveyors of falsehood beware:
I’m sick of your lies, and, truth be told,
I am the truth, behold!

The above account of truth is clearly satirical since no one would seriously grant that the truth of all statements is grounded in the assertions of one individual person. But what are the more serious alternatives for definitions of truth? As usual in philosophy, there’s much disagreement about what the correct definition is. We will consider the three leading candidates here.

The first and most famous definition of truth is the correspondence theory: a statement is true if it corresponds to fact or reality. This is the most commonsensical way of looking at the notion of truth and is how standard dictionaries define the concept. A true statement simply reflects the way things really are. Take the statement “My car is white.” This statement is true if it conforms to how the world actually is, specifically whether my car is in fact painted white. As compelling as the correspondence theory of truth seems, skeptics immediately see one major flaw with it: we don’t have access to the world of facts. In spite of my best efforts to discover the way things really are, I’m at the mercy of my five senses, which, we’ve seen, are unreliable. While my senses tell me that my car is white, the color receptors in my eyes may not be working properly and my car may be a shade of yellow. For that matter, I may be living in a world of hallucinations and don’t even own a car. The sad fact is I can never reach beyond my perceptions and see the world as it really is.

With trivial issues, such as the truth concerning the color of my car, I may be willing to simply pretend that I have direct access to the world of facts and blindly trust my senses. This may serve my immediate needs perfectly well. It isn’t so easy to pretend, though, when I investigate the truth of more serious statements, such as whether “Bill murdered Charlie.” Even if I have a mountain of evidence that implicates Bill, such as fingerprints and eyewitness testimony, it’s impossible for me to turn back the hands of time and directly access the scene of Charlie’s murder. I only have hints about what the reality is. Similarly, if I’m investigating the truth of the statement “God exists,” I can’t directly access the reality of an infinitely powerful deity, even if God did exist and stood right in front of me. The best I would have is some imperfect evidence that the mysterious being standing before me was indeed God. Thus, the correspondence theory would not permit us to say either that “It is true that Bill murdered Charlie” or “It is true that God exists.”

A second famous definition of truth is the coherence theory, which aims to address the shortcomings of the correspondence theory.
According to the coherence theory, a statement is true if it coheres with a larger set of beliefs. Rather than attempting to match up our statements with the world of facts, we instead try to see if our statements mesh with a larger web of beliefs that support them. For example, the statement “my car is white” is true if it coheres with a collection of other beliefs such as “many cars are painted white,” “I perceive that my car is white,” and “other people invariably report that my car is white.” With the coherence theory, we avoid skeptical obstacles such as the unreliability of our senses and the possibility that we are hallucinating. What matters is our web of beliefs, which we all have access to— in contrast with a hidden world of facts that is blurred by the limits of our sensory perceptions. We also can even investigate statements such as “It is true that Bill murdered Charlie” or “It is true that God exists.” What matters here is whether these statements consistently fit with other beliefs that we have—beliefs about the pieces of evidence against Bill and beliefs about the evidence regarding a divine being.

Unfortunately the coherence theory faces serious criticisms, the most important of which is that it is relativistic. That is, it grounds truth in the changeable beliefs of human beings, rather than in an unchanging external reality. According to the coherence theory, the standard for all truth is the larger web of beliefs that people hold—beliefs about white cars, criminal evidence, evidence for God’s existence, and countless other issues. The problem is that belief systems come and go. Take beliefs about criminal evidence as just one example. Many cultures throughout history based criminal convictions on the evidence of supernatural omens: prophetic visions, the flight path of birds, patterns in the guts of sacrificed animals. That was their belief system which they relied on. In other cultures the testimony of one eye witness was sufficient to prove guilt. In our culture today we have fingerprints, DNA samples and psychological profiles which all contribute to our belief system about criminal guilt. The statement “Bill murdered Charlie” could cohere with some belief systems, but not with others. We typically think about truth as being absolute: either Bill murdered Charlie or he didn’t. If truth hinges on a changeable belief system, though, truth is no longer absolute.

The problems with the correspondence and coherence theories are so serious that many contemporary philosophers have abandoned both. In fact some philosophers have even abandoned the concept of “truth” as being completely unnecessary. This brings us to our third theory, the deflationary theory of truth: to assert that a statement is true is just to assert the statement itself. Compare these two statements:

- My car is white.
- It is true that my car is white.

What is the difference between the two? Nothing of substance. The phrase “it is true that” seems to be just repeating something that is already assumed in the phrase “my car is white.” In that sense, I am being redundant if I use the phrase “it is true that.” At times it may be rhetorically helpful to use the phrase “it is true that” in an effort to convince someone of my belief. Suppose you say to me “I don’t believe that your car is white.” I might respond by saying, “You’re wrong: it’s absolutely true that my car is white.” Again, I’ve not added anything of substance by injecting the notion of truth into my response; I’ve just stood up to you more forcefully. In short, according to the deflationary theory, the quest for a clear conception of truth—such as correspondence or coherence—will not succeed because it is ultimately a quest for something that doesn’t really exist.

But the deflationary theory also faces problems, one of which is that the notion of truth is built into our normal expectations of what we assert. When I say that “my car is white” you have an expectation that what I’m saying is true. Occasionally I do say
something that is false, but when that happens we all recognize that I’m doing something that is incorrect. The normal expectation, then, is that my assertion will be truthful. And this creates a problem for the deflationary theory: by eliminating the notion of truth, it cannot adequately account for our normal expectation of truthfulness.

Theories of Justification. Of the three components of knowledge, justification is the one that has attracted the most attention among contemporary philosophers. For centuries most philosophers followed a theory of justification called foundationalism. On this view, our justified beliefs are arranged like bricks in a wall, with the lower ones supporting the upper ones. These lowest bricks are called “basic beliefs,” and the ones they support are “non-basic” beliefs. Take this example:

- My car is white (non-basic belief)

This belief rests upon some supporting ground-level basic beliefs, including these:

- I recognize the car in front of me as my car (basic belief)
- I remember what white things look like (basic belief)
- The car in front of me looks white (basic belief)

There are two distinct elements to this foundationalist theory of justification. First, our ground-level basic beliefs are self-evident, or self-justifying, and thus require no further justification. When we have such beliefs, we cannot be mistaken about them, we cannot doubt them, and we cannot be corrected in our beliefs about them. For example, if I am perceiving the color white, then my belief that I am perceiving white is self-evident in this way. Even if I am hallucinating at the moment, my belief that I am perceiving the color white cannot be called into question. The second element of foundationalism is that justification transfers up from my foundational basic beliefs to those non-basic beliefs that rest upon them. Think of it like the mortar between bricks that begins at the very bottom level, locks them solid, and moves upwards to lock the higher bricks into place. For example, if I have the three basic beliefs about my car and whiteness listed above, then I am justified in inferring the non-basic belief that “my car is white.”

While foundationalism holds a respected place in the history of philosophy, it faces a major problem: it is not clear that there really are any self-evident basic beliefs that form the foundation of other beliefs. Foundationalists themselves have mixed views about what exactly our lowest-level foundational beliefs are. Descartes, for example, argued that there is only one single brick at the foundation of my wall of beliefs, namely, my belief that I exist. Every other belief I have rests on this. Locke, on the other hand, held that our most foundational beliefs are simple perceptions such as blue, round, sweet, smooth, pleasure, motion. These combine together to make more complex ideas. Contemporary philosophers resist both Descartes’ and Locke’s depiction of our most foundational beliefs. Some offer examples such as “I see a rock” (a basic belief about one’s perception), “I ate cornflakes this morning” (a basic belief about one’s memory), or “That person is happy” (a basic belief about another person’s mental state). But even these are questionable since they seem to rely on beliefs or perceptions that are more ground-level. If there really are ground-level foundational beliefs that are self-evident or self-justifying, you’d think that philosophers would have agreed along time ago about exactly which ones they are. But there is no such agreement.

An alternative to foundationalism is coherentism: justification is structured like a web where the strength of any given area depends on the strength of the surrounding areas. Thus, my belief that my car is white is justified by a web of related beliefs, such as these:
• I recognize the car in front of me as my car
• I remember what white things look like
• The car in front of me looks white

These, though are not foundational, but instead depend on another web of beliefs related to them, which includes these:

• I remember purchasing my car
• People seem to agree that I use the term “white” properly
• Nothing is abnormally coloring my vision, such as a pair of sun glasses

Each of these, in turn, rests on an ever-widening web of related beliefs. At no point do we reach a bottom-level foundation to these beliefs; the justification of each belief rests on the support it receives from the surrounding web of beliefs that relates to it. Coherentism is closely associated with the coherence theory of truth. With truth we determine that a proposition is true if it coheres with a larger web of beliefs. With justification, we determine that a belief is justified if it is supported by a larger web of beliefs. Coherentism’s similarities with the coherence theory of truth make it vulnerable to the same fundamental charge of relativism: not everyone’s belief system is the same, so a particular belief might find justification within your larger web of beliefs, but not within mine. Your belief system might justify the belief that “Bill killed Charlie,” that “God exists,” or that “abortion is immoral,” while my belief system might not justify any of these. We’d like to think that justification is a bit more universal and not dependent on the peculiarities of a particular person’s belief system.

Given the liabilities of both foundationalism and coherentism, many contemporary philosophers hold a third position called \textit{reliabilism}: justified beliefs are those that are the result of a reliable process, such as a reliable memory process or a reliable perception process. It’s like how we depend on a reliable clock to tell us what time it is. As long as we have confidence in the clock mechanism itself, then that’s all we need in order to trust the time that it tells us. We don’t have to inspect the internal gears of the clock and see how they relate to the movement of the clock’s hands. Similarly, to justify my beliefs, I don’t need to inspect how each belief connects with surrounding beliefs that are beneath them or next to them; I just trust the reliability of my mental process that gives me the belief. If my memory process is on the whole reliable, then I’m justified in my belief that I ate cornflakes this morning for breakfast. If my perceptual process is on the whole reliable, then I’m justified in my belief that my car is white. What matters is the reliability of the larger processes upon which my beliefs rest, not my other beliefs that border them.

According to reliabilism, the fault with both foundationalism and coherentism is that they rely too much on introspection: presumably, with our mind’s eye, we can see the strength of our specific beliefs and how they gain support from other beliefs that are connected to them (either like bricks in a wall or strands in a web). But, says the reliabilist, this approach places too much confidence in our ability to internally witness the connections between our specific beliefs. Our standard of justification should not depend on what our mysterious mind’s eye internally perceives, but, instead, upon more external standards and mental processes that we know are reliable through our life experiences. I am justified in believing that I ate cornflakes for breakfast because that’s what I remember, and I trust my memory since it is a reliable process of supplying me with information about the past.

\textbf{What’s So Bad about Relativism?} Twice so far the issue of relativism has raised its ugly head. The relativist position in general is that knowledge is always dependent upon some particular
conceptual framework (that is, a web of beliefs), and that framework is not uniquely privileged over rival frameworks. The most famous classical statement of relativism was articulated by the Greek philosopher Protagoras (c. 490–c. 420 BCE), who said that “Man is the measure of all things.” His point was that human beings are the standard of all truths, and it’s a futile task to search for fixed standards of knowledge beyond our various and ever-flexible conceptual frameworks.

Knowledge in medieval England depended on the conceptual framework of that place and time. Knowledge for us today depends on our specific conceptual frameworks throughout the world and throughout our wide variety of social environments.

Our initial reactions to relativism are usually negative. “The truth is the truth,” I might say, “and it shouldn’t make any difference what my individual conceptual framework is. Some conceptual frameworks are simply wrong, and others may be a little closer to the truth.” But is relativism really so bad that it warrants this negative reaction?

The first step to answering this question is to recognize that there are different types of relativism, some of which may be less sinister than others. The most innocent and universally accepted type is **etiquette relativism**, the view that correct standards of protocol and good manners depend on one’s culture. When I meet people for the first time, should I bow to them or shake their hand? If I make the wrong decision, I might offend that person, rather than befriend them. Clearly, that depends on the social environment that you’re in, and it makes no sense to seek for an absolute standard that applies in all situations. Etiquette by its very nature is relative.

There is also little controversy regarding **aesthetic relativism**, the view that artistic judgments depend on the conceptual framework of the viewer. We commonly feel that there is no absolute right and wrong when it comes to art, and it’s largely a matter of opinion. I might enjoy velvet paintings of dogs playing cards, while that might offend your aesthetic sensibilities. In many cases, **perceptual relativism** is also no big issue: one’s sensory perceptions depend on the perceiver. Something might appear red to me but green to you. There are people known as “supertasters” who experience flavors with far greater intensity than the average person, so much so that they need to restrict themselves to food that you or I would find bland. How we perceive sensations depends on our physiology, which we readily acknowledge may differ from person to person.

The types of relativism that we often resist, though, are those connected specifically with the two components of knowledge that we’ve discussed above, namely, truth and justification. **Truth relativism** is the view that truth depends upon one’s conceptual framework. This amounts to a denial of the correspondence theory of truth and acknowledges our inability to access an objective and independent reality. **Justification relativism** is the view that what counts as evidence for our beliefs depends upon one’s conceptual framework. This is a denial of foundationalism and an acknowledgement of coherentism. The German philosopher Friedrich Nietzsche (1844-1900) boldly embraced truth and justification relativism, as we see here:

> Positivism stops at phenomena and says, “These are only facts and nothing more.” In opposition to this I would say: No, facts are precisely what is lacking, all that exists consists of interpretations. We cannot establish any fact “in itself”: it may even be nonsense to desire to do such a thing. . . . To the extent to which knowledge has any sense at all, the world is knowable: but it may be interpreted differently, it has not one sense behind it, but hundreds of senses. “Perspectivity.” [Will to Power, 481]

For Nietzsche, then, there are many perspectives from which the world can be interpreted when we make judgments. Some
justification relativists even go so far as to deny the universal nature of so-called laws of logic; even these, they maintain, are grounded in mere social conventions.

A standard criticism of truth and justification relativism is that it leads to absurd consequences that no rational person would accept. By surrendering to relativism, we abandon any stable notion of reality and place ourselves at the mercy of cultural biases, fanatical social groups, and power hungry tyrants who are more than happy to twist our conceptual frameworks to their benefits. Everything, then, becomes a matter of customs that are imposed on us, even in matters of science. Scottish philosopher James Beattie (1768–1790) makes this point in a fictional story where he describes a crazy scientist who attempts to put relativism into practice:

[The scientist] was watching a hencoop full of chickens, and feeding them with various kinds of food, in order, as he told me “that they might [give birth to live offspring and] … lay no more eggs,” which seemed to him to be a very bad custom. . . . “I have also,” continued he, “under my care some young children, whom I am teaching to believe that two and two are equal to six, and a whole less than one of its parts; that ingratitude is a virtue, and honesty a vice; that a rose is one of the ugliest, and a toad one of the most beautiful objects in nature.” [James Beattie, “The Castle of Skepticism”]

According to Beattie, if we took the relativist’s position seriously, we’d be forced to accept absurd views like “it is just a matter of custom that chickens lay eggs,” or that “it’s possible that 2+2=6.” Thus, even if we acknowledge a certain level of relativism with etiquette, aesthetics and perception, we need to draw the line when it comes to standards of truth and justification.

How might the relativist respond to this criticism? One approach is to hold that not all conceptual schemes are on equal footing, and some indeed are better than others. Nietzsche argues that there are competing perspectives of the world, and the winner is the one whose conceptual framework succeeds the best:

It is our needs that interpret the world; our instincts and their impulses for and against. Every instinct is a sort of thirst for power; each has its point of view, which it would gladly impose upon all the other instincts as their norm. [Will to Power, 481]

Nietzsche presents the conflict as a kind of power struggle among competing conceptual frameworks, where the winner takes all. A more gentle approach, though, would be to hold that the winner is the one that best assists us in our life’s activities and allows us to thrive. If people today held that “it is just a matter of custom that chickens lay eggs,” or “it is possible that 2+2=6,” their underlying conceptual framework would not enable them to succeed very well in the world. For that matter, such a conceptual framework would not have allowed people to thrive very well in medieval England or any other pre-modern period of human history. While there may be an underlying objective reality that molds our conceptual frameworks in successful ways, that possibility is irrelevant since, according to the relativist, we could never know such an objective reality even if it existed. What we do know is how our conceptual frameworks enable us to succeed in the world, and that’s the real litmus test for truth and justification.

Thus, with many ordinary life beliefs, relativist theories of truth and justification work reasonably well, without leading us down the path to absurd consequences. What, though, of more scientific theories? In medieval times people thought mental illness was caused by demon possession; today we think that it is caused by physiological brain disorders. The medieval theory worked well in its own day; does that mean that it was true back then—supported
by its own web of beliefs—but not now? In scientific matters, people feel uncomfortable with relativism and instead believe that our knowledge of physics, chemistry and biology has a fixed and objective reference point. We will next examine the issue of scientific knowledge in more detail.

E. Scientific Knowledge

Every child knows the tale of Isaac Newton’s inspiration for his views on gravity: while sitting beneath a tree he saw an apple fall, which prompted him to wonder why things always fell downward rather than sideways or upward. In time Newton formulated his theory of universal gravitation, which described the attraction between massive bodies. Less known is the rival theory of intelligent falling, recently devised by the satirical newspaper *The Onion*. According to this view, things fall downward “because a higher intelligence, ‘God’ if you will, is pushing them down.” As proof for their view they cite a passage from the Old Testament book of Job: “But mankind is born to trouble, as surely as sparks fly upwards.” Accordingly, a defender of intelligent falling theory remarks, “If gravity is pulling everything down, why do the sparks fly upwards with great surety? This clearly indicates that a conscious intelligence governs all falling.” The theory of intelligent falling is obviously not a real theory, but rather a parody of the religiously-based intelligent design theory. Nevertheless, we can ask the basic theoretical question, why is universal gravitation a better account of natural events than intelligent falling? The job of science is to explain how the natural world works, to give us knowledge of the underlying mechanics of natural phenomena. That knowledge does not come easy, though, and it seems that science has to wrench nature’s secrets out of her. As scientists put forward rival theories, how do we determine which are closer to the truth?

**Confirming Theories.** The starting point for discussion is to distinguish between three related scientific concepts: a hypothesis, a theory, and a law. The weakest of these is the *scientific hypothesis*, which is any proposed explanation of a natural event. It is a provisional notion whose worth requires evaluation. Newton’s account of gravity began as a humble hypothesis, and even the theory of intelligent falling qualifies as a hypothesis. While hypotheses may be inspired by natural observations, they don’t need to be, and virtually anything goes at this level. One step up from this is a *scientific theory*, which is a well confirmed hypothesis. It is not mere guess, like a hypothesis may be, but is a contention supported by experimental evidence. When Newton proposed his account of gravity, he accompanied it with a wealth of observational evidence, which quickly elevated it to the status of a theory. This, though, is where the theories of gravity and intelligent falling part company: there’s no scientific evidence in support of intelligent falling, and thus it fails as a theory. Lastly, there is a *scientific law* which is a theory that has a great amount of evidence in its support. Indeed, laws are confirmed by such a strong history of evidence that they cannot be overturned by any singular piece of evidence to the contrary; rather, we assume instead that that singular piece of contrary evidence is flawed. As compelling as Newton’s theory of gravity was, it took well over 100 years before it was confirmed to the point that it gained status as a law.

We see that confirmation is the critical component in establishing a scientific claim: it is what elevates a hypothesis to a theory, and a theory to a law. There are several different ways of confirming scientific notions. The first factor in the confirmation process is that it exhibit *simplicity*; that is, when evaluating two rival theories, the simplest theory is the one most likely to be true. This doesn’t guarantee that it’s true, but, all things being equal, it’s the one that we should prefer. Compare, for example, universal gravity and intelligent falling. Universal gravity involves a single gravitational force that is inherent to all physical bodies. Intelligent falling, on
the other hand, involves countless divine actions that guide individual bodies downwards. We should thus prefer universal gravity as the correct explanation since it is not burdened by such an abundance of distinct divine actions.

A second component of confirmation is unification, that is, the ability to explain a wide range of phenomena. The rule of thumb here is that the more information explained by a theory, the better. Science is an immense interrelated system of facts, laws, and theories, and scientific contentions gain extra weight when they contribute to the scheme of unification. It is unification that gave an initial boost to Newton’s theory of universal gravitation. Prior to Newton, Astronomers assumed that planets and other celestial objects followed their own unique set of laws that were distinct from those on earth. However, Newton showed how the motions of the planets were governed by precisely the same rules of gravity and motion that physical bodies on earth obey.

A third factor in scientific confirmation is successful prediction. Good scientific theories should not simply organize collections of facts, but should be able to reach out and predict new phenomena. This is what bumped Newton’s theory of gravity up to the status of a law. Astronomers in the early 19th century noticed some strange movements in the orbital pattern of the planet Uranus, and they hypothesized that the irregularities were caused by the gravitational tugging of an undiscovered eighth planet. Applying Newton’s formulas of gravity and motion, they pinpointed a location in space where the large body must be. Then, pointing their telescopes at the spot, they discovered the mystery planet, which was subsequently named “Neptune.” Scientific predictions like these don’t happen too often, but when they do they do much to confirm a theory. Einstein’s theory of relativity, for example, was confirmed with the prediction of bent star light during a solar eclipse.

A fourth and final factor in scientific confirmation is falsifiability: it must be theoretically possible for a scientific claim to be shown false by an observation or a physical experiment. This doesn’t mean that the scientific claim is actually false, but only that it is capable of being disproved. The criterion of falsifiability is important for distinguishing between genuine scientific claims that rest on tests and experimentation, and pseudo-scientific claims that are completely disconnected with testing. Take, for example, the views of Heaven’s Gate believers that we examined at the outset of this chapter. According to them, aliens come down to earth in the form of teachers, but once becoming human they are stripped of their previous memories and knowledge. Thus, we can’t test the claims of these teachers about their previous alien lives, since they can’t remember anything about them. “So tell me a little about your home planet” I might ask one alleged alien. He then replies “Sorry, I can’t remember anything about it, but I’m still an Alien.” To make things worse, Heaven’s Gate believers claim that the aliens purposefully imposed this knowledge restriction on themselves since “too much knowledge too soon could potentially be an interference and liability to their plan.” In short, their claims about the aliens are completely resistant to refutation. Fortune tellers are another good example of this, as the philosopher Karl Popper explains here:

By making their interpretations and prophesies sufficiently vague they were able to explain away anything that might have been a refutation of the theory had the theory and the prophesies been more precise. In order to escape falsification they destroyed the testability of their theory. It is a typical soothsayer’s trick to predict things so vaguely that the predictions can hardly fail: that they become irrefutable. [Conjectures and Refutations]

Legitimate scientific theories, by contrast, always hold open the possibility of being refuted by new data or a new experiment. By
putting forth their theories, scientists take a risk that what they’re proposing might be disproved by the facts. Even universal gravitation is vulnerable to refutation if some future experiments produce compelling evidence against it. Thus, a good theory is always potentially falsifiable, although it has not been actually falsified.

**Scientific Revolutions.** Scientists continually push the boundaries of knowledge, and on a daily basis we see new theories about the spread of diseases, healthy eating habits, or the environment. We also read about new studies that challenge previously accepted scientific views. For example, in contrast to earlier claims by scientists, the accepted wisdom now is that vitamin C does not help prevent colds, and fiber in our diets does not help prevent colon cancer. Science thus moves ahead in baby steps, occasionally taking a step backwards to correct an erroneous theory. All the while, though, the larger body of scientific knowledge seems secure and well established. But then sometimes a new scientific theory comes along that is so radical and far reaching in its consequences that it forces scientists to throw out many of their underlying assumptions about the world and set things on a dramatically new course. These are *scientific revolutions.* The most dramatic example is the shift from the earth-centered view of the heavens, championed by the ancient Greek astronomer Ptolemy, to a sun-centered system which was defended by Copernicus in the 1500s. This “Copernican Revolution,” as it is now called, did more than simply swap the earth with the sun in the model of celestial objects. It also had the effect of overthrowing medieval theories about matter and motion, and ultimately replacing them with Newton’s laws of motion. Other important revolutions were sparked by Charles Darwin’s account of evolution, Einstein’s account of general relativity, and the Big Bang theory.

The most probing philosophical analysis of scientific revolutions was offered by American historian of science **Thomas Kuhn** (1922-1996). Kuhn argued that scientific revolutions are the result of changing *paradigms*—that is, the web of scientific beliefs held in common by members of the scientific community. When major paradigms are overturned and replaced with new ones, such as replacing the Ptolemaic with the Copernican paradigm, we have a scientific revolution. What triggers the paradigm shift, according to Kuhn, is that scientists run into inconsistencies with the old paradigm that cannot easily be explained away. Scientific theories will always face some irregularities—such as with an experiment that seems to contradict an accepted theory. If the theory is well established, a few irregularities here or there won’t matter; in fact, scientists often chalk these up to an acceptable level of error that’s built into the enterprise of scientific investigation. However, sometimes irregularities pile up to such a degree that it throws science into a condition of crisis. Seeking resolution to the crisis, scientists then replace an old scientific paradigm with a new one that better resolves the irregularities.

Kuhn argues that scientific revolutions have much in common with political revolutions: rebel groups think that the ruling institution is inadequate, which they then overthrow and replace with a new one:

Political revolutions are inaugurated by a growing sense, often restricted to a segment of the political community, that existing institutions have ceased adequately to meet the problems posed by an environment that they have in part created. In much the same way, scientific revolutions are inaugurated by a growing sense, again often restricted to a narrow subdivision of the scientific community, that an existing paradigm has ceased to function adequately in the exploration of an aspect of nature to which that paradigm itself had previously led the way. In both political and scientific development the sense of malfunction that can lead to crisis is prerequisite to revolution.
Kuhn warns that the transition from the old to new paradigm is not a smooth one. Many scientists will hold fast to the old paradigm, and the new paradigm needs to attract an ever-growing number of supporters before it can finally overthrow the old one. The upshot of Kuhn’s theory is that science is not cumulative: our present theories are not built upon the secure foundation of past theories. Instead, scientific knowledge shifts according to our current paradigms, and, once again, the issue of relativism arises. These paradigms are webs of belief that are held by the scientific community at the time. Truths in science, then, are relative to these shifting webs of belief.

Kuhn’s account of scientific revolutions has its critics, particularly among those who believe that science, when done properly, is grounded in objective truth, and not in shifting belief paradigms of the scientific community. One criticism is that Kuhn has over dramatized the sweeping nature of most scientific revolutions. Sure, the Copernican revolution was indeed a major one that resulted in overthrowing old scientific models that were rooted in superstitious conceptions of the world and sloppy experimentation. In fact, the older models were so engrained with religious mythology and metaphysics, it’s overly generous to call them “scientific.” Since the time of Copernicus, however, we’ve not seen any scientific revolutions that “overthrow” entire paradigms. Rather, new mini-revolutions seek to encompass much of the theory and data of previous scientific investigations while at the same time setting a new direction for future investigation. For example, Newton’s laws of motion were not overthrown by Einstein’s theory of relativity; instead scientists try to incorporate both into a larger scientific vision of reality that unifies all of nature’s forces. Other mini-revolutions, such as Darwinian evolution, quickly put an end to rival theories of biological development, such as Lamarck’s, that had little or no supporting evidence to begin with. Thus, contrary to Kuhn’s position, when science is done properly our knowledge of the natural world is cemented into a fixed and objective reference point.

This chapter began with a discussion of radical skepticism, and while that may not be the most cheery way of investigating the nature of knowledge, in many ways it sets the right tone. No matter what we say to clarify the characteristics of knowledge, warning flags immediately go up. All of our sources of knowledge have serious limitations. The very definition of knowledge can be picked apart by an endless variety of Gettier-type counter examples. Theories of truth and justification seem to be either naively optimistic, or they lean towards relativism. While scientific knowledge attempts to move progressively towards unchanging truth, it is always cradled by a potentially changing web of beliefs held by scientists. Achieving genuine knowledge is in some ways like playing a video game where the winning score is infinitely high: no matter how close we move towards it, it remains at a distance. If the human effort to gain knowledge was merely a leisure activity like playing an impossible game, we’d certainly give it up for a more attainable diversion. But the pursuit of knowledge is a matter of human survival that we can’t casually set aside. Philosophical discussions of knowledge are an important reality check as we routinely gather facts and construct theories about how the world operates. The hope of acquiring a fixed body of knowledge is very seductive, and the problems of knowledge that we’ve covered in this chapter help us resist that temptation.

For Review

1. What are the three main arguments for radical skepticism?
2. What are the four main criticisms of radical skepticism?
3. What are the four main sources of experiential knowledge?
4. What are the key features of non-experiential knowledge?
5. What are the key features of rationalism and empiricism?
6. Describe the “JTB” definition of knowledge.
7. What is the point of the Gettier problem?
8. Name and describe the different theories of truth.
9. Name and describe the different theories of justification.
10. Name and describe the different types of relativism.
11. What are the different ways in which scientific theories and laws gain confirmation?
12. What is Kuhn’s view of scientific revolutions?

For Analysis

1. Write a dialogue between a radical skeptic who thinks that he’s living in an artificial reality, and a non-skeptic who thinks we have knowledge of the commonsense world that we perceive.
2. Explain the different features of rationalism and empiricism, and try to defend one position over the other.
3. Explain the foundationalist theory of justification, and try to defend it against one of the criticisms.
4. Write a dialogue between a relativist respond and non-relativist regarding the argument against relativism from absurd consequences.
5. Explain the notion of falsification, and then describe whether a religious view like creationism, intelligent design theory, or intelligent falling theory can be falsified.
6. Explain the criticism of Kuhn at the end of the chapter and try to defend Kuhn against it.

REFERENCES AND FURTHER READING

Works Cited in Order of Appearance

Information about the Heaven’s Gate cult can be found at http://www.rickross.com/groups/heavensgate.html


Further Reading


CHAPTER 6: ETHICS

A. Moral Objectivism and Relativism
   Moral Objectivism
   Moral Relativism
   The Moderate Compromise

B. Selfishness
   The Case for Egoism
   The Case for Altruism
   Egoism and the Struggle for Survival

C. Reason and Emotion
   Moral Reasoning: Detecting Truth and Motivating Behavior
   We Can’t Derive Ought from Is
   Moral Utterances Express Feelings

D. Virtues
   The Virtuous Mean
   Virtues and Gender
   Virtues and Rules

E. Duties
   Duties to God, Oneself and Others
   The Categorical Imperative
   Duties to Animals and the Environment

F. Utilitarianism
   The Utilitarian Calculus
   Higher Pleasures and Rules
   Reactions from Duty Theorists

For Reflection

1. What are some strange moral practices in other cultures that you’ve heard about?
2. Suppose that I rescue someone from drowning. What might be a selfish motive behind my action?
3. Imagine that you were debating the issue of abortion with someone and your conversation became heated. What emotionally-charged things might you say to each other on this topic?
4. Imagine that your beloved uncle died, and at the viewing everyone is talking about what a wonderful person he was. Someone, for example, says that your uncle was a very generous man. What other flattering character traits might the mourners describe?
5. Are there any moral principles that people know instinctively? What about “Don’t steal”?
6. Can an action be morally right if it produces more bad consequences than good ones? Try thinking of an example.

One February morning two armed gunmen wearing black ski masks entered a Los Angeles bank, fired their machine guns and ordered everyone to the ground. Bank personnel gave them $300,000, but when leaving they were met by police who had already surrounded the building. Undaunted, the two men stepped into the street, firing at the officers. The police responded in kind, but their bullets bounced off the robbers who were covered head to toe with body armor. More officers were called in as the robbers casually walked down the road, shooting at everything in sight. Soon 200 police were on the scene, but even then they were so overpowered by the robbers that they had to borrow additional guns and ammunition from a local gun store. The robbers fired over 1,200 bullets during the 45 minute battle. Eventually, one gunman, who was shot 11 times, ended his life with a bullet to his head. The other surrendered after being shot 29 times, but he died five minutes later from blood loss. Miraculously, no one else died, although many were injured.

In a bizarre postscript to this story, three years later family members of the second gunman took the police to court, claiming that officers intentionally let him bleed to death. During the trial, the lawyer for the gunman’s family conceded that the robber “did a
very bad thing. He robbed a bank and shot a lot of people.” Nevertheless, the attorney argued, the police had a responsibility to get him medical treatment, which they failed to do. In a vote of 9 to 3, the jury found that the gunman’s rights were not violated.

This story presents us with an intricate web of ethical concepts: bad behavior, devastating consequences, moral outrage, rights, responsibilities, selfishness and greed. It’s one thing for us to intuitively feel that the gunmen were immoral, but it’s another to say clearly what that immorality consisted of. For 2,500 years philosophers have been trying to unravel the mysteries of ethical judgments, and their theories differ dramatically. In this chapter we will look at some of the more famous of these theories.

A. Moral Objectivism and Relativism

An initial puzzle about morality concerns its ultimate source: where does morality come from? Do we create it or is it etched into the stars? Imagine that in a distant country two bank robbers in body armor blasted off 1,200 bullets; when they finally surrendered the Mayor gave the gunmen the Keys to the City in reward for their outstanding display of strength and courage. How would you respond to this story? On the one hand you might say that the Mayor was crazy: what the gunmen did was simply wrong and their actions should be punished, not rewarded. There’s an unchanging standard of justice that everyone must abide by, wherever they are in the world. This is the position of moral objectivism, which, at the risk of oversimplification, has three main features that many traditional philosophers embrace. First, morality is objective in the sense that ultimate moral standards exist independently of humans. Often this is described as a spiritual realm which is fixed and permanent, unlike the physical world of human society that is in constant flux. Second, moral standards don’t change throughout time; they are eternal. The true standards of morality that people followed thousands of years ago still apply today and will apply in the future. Third, moral standards are universal in the sense that they apply to everyone. No one can claim to be immune from the demands that true morality places on us all.

On the other hand, upon hearing this revised story about the gunmen, you might insist that the Mayor did nothing wrong by rewarding them: that’s just how they do things in that country. We have one concept of justice, they have another; neither way is superior, they’re just different. This is the position of moral relativism, which also has three key ingredients. First, morality is a purely human invention. People create morality, and it by no means exists independently of humans, such as in some higher spiritual realm. Second, moral standards change throughout time and from country to country [and perhaps even person to person]. Wherever we go in the world we will find radically different moral values. Third, moral standards do not apply universally to all people; it is instead relative to our unique situations.

Moral Objectivism. Moral objectivism and moral relativism each have long and distinguished histories. The most important proponent of objectivism was Plato (428–348 BCE), who developed a two-tiered picture of the universe. The lower level consists of physical things like rocks, trees, houses and physical human bodies. Things at this level are a major disappointment; they shift and change, they cause misery, and they distract us from genuine truth. The higher level is spiritual in nature, and consists of unchanging truths, particularly those of mathematics and morality. When we behave morally, we are in fact molding our actions after the higher spiritual standard of justice, that is, after the forms, as Plato calls them. When acting justly, for example, we are in fact setting aside human conventions and grasping the ultimate form of justice. In addition to justice, there are forms of charity, courage, wisdom, and perfect goodness.
According to Plato, when relativists claim that morality is an ever-changing human creation, they are simply blinded by the corrupt physical world and incapable of mentally perceiving the forms. Relativists are not entirely to be blamed, he argues, because spiritual vision is a difficult thing to acquire. We need to be awakened to the flaws in the world around us, and use a special and often hidden mental capacity to access the spiritual realm. As radical as Plato’s view of morality is, it was nevertheless embraced by philosophers throughout history, and it only declined in popularity in the past 200 years. Its special appeal is its emphasis on the universal nature of morality. All of us at one point or another in our lives have felt that some moral principles apply to everyone. What Hitler did to the Jews, for example, was simply wrong, and no justification can be made on the grounds of the German culture of that time.

While Plato’s theory of the forms is one of the more philosophically sophisticated accounts of moral objectivism, there are other suggested ways of grounding morality in an objective reality. One of these is religious: morality is a creation of God’s will and is commanded by him for all of us to perform. God’s very nature, then, is the source and substance of morality; in a sense, the objective nature of morality is God’s mind. While believers in many religious traditions have embraced this theory, it has an important obstacle: even if God does exist, it’s not clear how we know what moral values he has commanded. Some suggest that God speaks to us directly, or that God has given us a moral conscience, or that God has inscribed moral rules into scripture. Each of these avenues of divine revelation, though, requires a great deal of trust. Should we believe someone who says that God told him to bomb an abortion clinic? Should we trust a political leader whose conscience tells him to wage a religious war against an enemy country? Should we trust a religious leader’s scriptural interpretation that some races are superior to others? These are extreme examples, but they suggest that claims to know God’s mind cannot be taken at face value. Rather, we often evaluate claims of divine revelation based on an independent standard of rightness that has nothing to do with God or religion.

**Moral Relativism.** Moral relativism is the offspring of philosophy’s long and controversial skeptical tradition. According to the Greek skeptic Sextus Empiricus (c. 160-210), if we want to achieve peace of mind, we should reject extreme notions of objective moral truth. Look around the world, he says, and you’ll find nothing but conflicting opinions about morality. One philosopher says that morality consists of pursuing pleasure, another says that it consists of resigning oneself to God’s will. Equally conflicting are people’s actual moral actions. One society eats their dead relatives, another feeds them to vultures. One society practices incest, another condemns it. He describes here conflicting attitudes about sexual morality:

We consider it shameful to have sex with a woman in public, but it is not thought so by some of the people from India. They had sex publicly with indifference, like the philosopher Crates, as the story goes. Additionally, prostitution is shameful and disgraceful among us, but for many of the Egyptians it is highly respected. They say that women with the greatest number of lovers wear an ornamental ankle ring as a proud token of their accomplishments. Some of the girls marry after collecting a dowry beforehand by means of prostitution. We find the Stoics maintaining that it is acceptable to keep company with a prostitute or to live on the profits of prostitution. [Outlines of Skepticism, 3:24]

According to Sextus, moral diversity is everywhere on virtually every issue. What, he asks, gives us the authority to prefer one person’s or society’s views over another? We’re in no position to make the judgment call. The safe thing to do is to doubt the
existence of objective morality, and see that all moral issues “are matters of convention and are relative.”

The central point here is the argument from cultural variation, which is this:

(1) If morality is objective, then we would not see widespread cultural variation in moral matters.

(2) There is in fact wide spread cultural variation in moral matters.

(3) Therefore, morality is not objective.

In Premise 1 the skeptic questions whether the notion of objective morality makes any sense, especially considering how people often behave in radically different ways. Objective morality may indeed be a pretty thought, but if it doesn’t reflect reality, then it is only a pretty thought. According to Premise 2, anthropological studies clearly show the enormous diversity in moral attitudes throughout the world. We regularly read about strange moral practices of foreign cultures, such as women who are stoned to death for having premarital sex. This makes us cringe on our side of the world, but people in the foreign country itself believe they are defending moral values that hold their society together. We have no choice, the skeptic concludes, but to reject moral objectivism and accept moral relativism as the only real alternative.

How might moral objectivists respond to this argument? Regarding Premise 1, according to objectivists, some people and societies are evil, and it’s no surprise that they concoct their own self-serving standards. Human nature is inherently flawed, perhaps because we are constructed out of flimsy material stuff, or perhaps because we are psychologically crippled. In either case, there is a barrier between us and the higher realm of moral truth, and weak people will wallow in their human-made filth. Regarding Premise 2, the objectivist continues, skeptics have greatly exaggerated the diversity that we see in moral matters. There is indeed great variation in cultural practices around the world, but how much of this is really moral? Many differences involve standards of mere taste or etiquette. For example, at great length Sextus describes strange funeral practices. There’s no question that these rituals are important in their respective societies, but they are matters of taste which cultures can rightfully determine for themselves. In their own ways, they all show an underlying respect for the deceased. Further, some values seem to be consistently endorsed by all cultures around the world, such as prohibitions against murder. It’s hard to take seriously any reports of societies that entirely lack this value. Thus, according to the objectivist, the entire argument from cultural variation falls flat.

The Moderate Compromise. The objectivist and relativist positions described above are rather extreme positions: either all moral values are unchanging, or all moral values are ever-changing human inventions. In spite of the enormous popularity of these extreme views, it seems that concessions are needed on both sides. Contrary to objectivists, cultural attitudes about some serious issues really do vary greatly, and it’s unreasonable to hunt for objective values that underlie these. For example, some societies feel that premarital sex and homosexuality are perfectly permissible; in other societies these are capital offences. However, contrary to relativists, some values appear consistently in human societies. This is so with prohibitions against murder and stealing. These values may not necessarily be grounded in the spirit-realm, but they are indeed universally endorsed within human societies—even though they are not always followed.

There are two ways that we might mediate between extreme objectivism and extreme relativism. The first assumes that there are at least some objectively-grounded moral standards, which are fixed and unchanging. These standards, though, are few in number and
very general in nature. Examples are that we should educate our children, avoid harming others, and help others in need. However, it is left to individual societies to interpret these abstract standards and derive more specific moral rules from these. For example, when attempting to clarify the moral obligation to educate our children, some societies may mandate compulsory public education as a way of teaching children; others may assign the task to parents in the home. Similarly, when interpreting the general moral standard that we should avoid harming others, societies might disagree as to what counts as “harm.” There is thus objectivity, permanence and universality with the most general moral principles. At the same time, though, there is relativity with the specific moral rules that we derive from these.

The second way of mediating between extreme objectivism and relativism is to emphasize the role of human social instincts. Assume for the moment that the skeptic is correct: there is no higher independent realm of morality and the concept is only a philosophical fable. Assume further that some key moral values are purely social creations and vary radically in different cultures—such as rules about sexual activity. Nevertheless, some moral values appear to be uniform, such as the need to avoid harming others and the need to show kindness and charity. These might be the result of social instincts that have formed in humans over thousands of years of biological evolution. We are social animals by our nature, and thus must have some instinctively-driven way of living in peace with each other. Some moral values then might be grounded in instinct, and so will appear uniform from person to person and throughout the world. Technically speaking, though, these values would not be eternal and unchanging: just as they emerged through evolution, they may just as quickly dispel through future evolutionary development. Nevertheless, for the time being there is some naturally-grounded uniformity that stands midway between the objectivist and relativist extremes.

Taken separately or together as a package, the above two compromises may help bridge the gap between our conflicting intuitions between objectivism and relativism. Ultimately, though, neither completely settles the issue, since at some level we’d still want to ask whether any moral values are completely permanent and unchanging. [To rely on the accident of having been raised in a certain religion or culture seems hardly enough to justify our moral revulsion at the sight of a mother beating her child in Wal-Mart or our admiration of an organ donor.]

B. Selfishness

When people rob banks, we presume that they are motivated through selfishness. In fact, much of our behavior throughout the day is motivated by self-oriented desires—to eat, sleep, relax, make money, impress other people. The interesting question, though, is whether all of our human actions are ultimately self-oriented. A postman was once delivering mail on his usual route when he saw smoke coming from a house. Knowing that an elderly man lived there, he knocked down the door and entered the house which, as he saw, had just caught fire from a kerosene heater. He dragged the old man towards the door, but both were overcome with smoke. Firefighters soon showed up and carried them to safety, although both were permanently injured. What motivates ordinary people to perform heroic deeds like this? One explanation is that people aren’t 100% selfish, and there is some instinctive capacity within human nature to help others irrespective of our private interests. That is, we have an instinct to act altruistically. But wait, the critic might say. Just because a postman rescues someone from a burning building, that doesn’t mean that he acted purely altruistically. Maybe he knew that he’d instantly become a hero, and he thought it was worth the risk to receive that honor. Maybe he thought that it was part of his postal job description, and he didn’t want to be reprimanded by his boss. His actions may have been motivated by entirely selfish reasons, regardless of how altruistic they appear on
the surface. That is, human nature may be entirely egoistic. The dispute is one of human psychology, and, more precisely, the competing principles are these:

*Psychological egoism*: human conduct is selfishly motivated, and we cannot perform actions from any other motive.

*Psychological altruism*: human beings are at least occasionally capable of acting selflessly.

The question is certainly an interesting one, but what does it have to do with morality? The answer is embedded in a basic moral principle: *ought implies can*. Decoded, this means that we are morally obligated to do only those things that we are capable of doing. It makes sense to say that I [unlike the kleptomaniac] am morally obligated to avoid shoplifting since that is within my power to do so. However, it does not make sense to say that I personally am responsible for curing cancer; lacking the required biochemical expertise, it’s not within my power to even attempt this, let alone accomplish this. Suppose, now, someone said that I have an obligation to donate to charity from completely altruistic motives—with no expectation of a tax break or any other benefit. Fulfilling that obligation depends in part on whether I’m psychologically capable of acting altruistically. This, though, is precisely what the egoist would deny. In short, if I am locked into behaving selfishly, then I have no moral responsibility to ever behave altruistically.

**The Case for Egoism.** One of the more notorious defenders of egoism was British philosopher **Thomas Hobbes** (1588–1679). Hobbes believed that human beings are biological machines that follow the same kind of physical laws that make clocks run. Our thought processes and actions are governed by our physiological makeup, and if we want to truly understand why we behave as we do, we must look to biology. He describes in detail how our various biologically-based emotions drive our conduct, and he makes clear that selfishness is the dominant motive behind our choices. Our instinctive selfishness emerges most dramatically, according to Hobbes, when life’s necessities such as food and shelter are in scarce supply and we quickly compete to acquire them before our rivals do.

When exploring the issue of egoism, Hobbes looks at two specific motives that we commonly think of as being altruistic: pity and charity. For example, suppose that, out of a genuine sense of pity and charity, the postman rescued the old man from the burning house. Ordinarily, we’d assume that these feelings are directed towards the victim, and not a reflection of the postman’s selfishness. However, according to Hobbes, even when we think such actions are altruistic, there is really a hidden selfish motivation. With any act of pity, we imagine *ourselves* in a position of distress. The postman thus imagined himself in the burning house and, based on that fictitious mental image, rescued the old man. With charity, Hobbes argued, we take special delight in exercising our power over other people. The postman thus wanted to feel the power of preserving the old man’s life, and so he attempted the rescue.

Although we may not agree with Hobbes’s specific psychological explanations of pity and charity, he nevertheless gives us a way of understanding so-called altruistic actions, namely, seeing them all as arising from hidden selfish motivations. Following Hobbes’s lead, we can construct a basic argument for psychological egoism here:

1. If we can adequately explain some phenomenon with one principle rather than two, then we should reject the second principle.
2. So-called altruistic behavior can be adequately explained through psychological egoism.
(3) Therefore we should reject the principle of psychological altruism.

Premise 1 is a basic principle of simplicity, which is sometimes called Ockham’s Razor, named after the medieval philosopher William of Ockham (c. 1288-1348) who regularly relied on it. Premise 2 maintains that we don’t really need the principle of instinctive altruism. This claim is particularly compelling since we already accept that selfishness is indeed a motivating principle of human nature. The question is whether we need to introduce another to explain so-called altruistic behavior. Hobbes says that we don’t.

The Case for Altruism. When Hobbes’s writings appeared, readers of the time were horrified by his defense of psychological egoism. In fact, Hobbes forced the issue and made it almost a requirement for moral philosophers after him to take a stand on the issue one way or another. One such philosopher was British clergyman Joseph Butler (1692-1752) who defended the principle of instinctive altruism against Hobbes’s attack. According to Butler, Hobbes’s fundamental error was oversimplifying the basic principles of human motivation. It looks neat and tidy to reduce everything to a single motive of selfishness as Hobbes did, but, Butler says, human nature is too complex to allow for this easy solution. Butler thus attacks Premise 2 in the above argument—while not disputing the principle of simplicity in Premise 1. He notes two specific errors of oversimplification in Hobbes’s theory.

First, according to Butler, Hobbes oversimplifies the notion of “selfishness”—or “self-love” as he also calls it. Yes, many of our actions are motivated specifically by self-love. However, some human inclinations might superficially appear to be the same as self-love, such as hunger and esteem, although they are actually different inclinations. Suppose, for example, that I am hungry and I eat a sandwich. Suppose that I even enjoy the experience of eating the sandwich. My motivation here is simply hunger and not self-love, since even if I hate myself I would still be motivated to eat and enjoy the sandwich. Similarly, with esteem, even if I hate myself, I could still desire to be valued by other people. So, according to Butler, Hobbes’s mistake was to reduce all self-oriented motives to the single theme of self-love. The egoist might accept Butler’s criticism and state more cautiously: all human action is motivated by a group of self-oriented motives. This revised version is still very much egoistic, and it denies instinctive altruism. However, by conceding this point, serious harm is done to the argument from simplicity. The main appeal of psychological egoism is its ability to offer a simplified account of human conduct. Now, though, we are not talking about a single selfish motivation, but a collection of self-oriented motives. At this stage, what would it hurt to toss in a couple of altruistic motives as well?

Second, Butler argues that we have an instinctive motive of benevolence that underlies human friendship, compassion, love, parental inclinations and other feelings. This becomes evident when we examine Hobbes’s egoistic interpretation of charity. If charity was motivated solely by our delight in exercising power over others, then how could we distinguish charity from sadistic cruelty which is precisely the same motive? The postman, for example, didn’t have to risk his own life to enjoy having power over the old man. He could have just stood there and watched, all the while delighting in the fact that he had the power to let the old man die. Obviously, says Butler, there’s something much more to charity than a power trip, and that extra something is instinctive benevolence. It’s simply not that easy to explain away human kindness.
Egoism and the Struggle for Survival. Evolutionary biologists from Charles Darwin (1809-1882) onwards have been captivated by the egoism–altruism debate. The theory of evolution is based on the assumption that organisms struggle to survive: those with the best survival skills live, and the rest eventually die. Selfishness is an integral part of survival, and it is difficult to find a place for self-sacrifice and kindness in the evolutionary struggle. The fact is, though, that people do sometimes behave kindly towards others—at least appearing to be altruistic—and this also requires an evolutionary explanation.

Evolutionary biologists largely reject the purest form of altruism, namely 100% selfless acts of kindness. Here’s why. Suppose that I and a fellow caveman are foraging through the woods looking for food. The pickings are slim, but at the same time we both spot a single apple on a tree. Suppose further that I have some instinct of genuine altruism in me, but the caveman is purely egoistic. So, moved by sympathy, I give the caveman my apple. Sadly, I starve to death and thus fail to pass my altruistic instinct on to my offspring. Meanwhile the egoist wines and dines a cavewomen with the apple, they mate, and he passes his egoistic traits onto their children. The moral of the story is that genuine altruism is not conducive to survival, and if any early humans ever had that trait, it would have been eliminated from the human gene pool long ago. But even if we reject genuine altruism from an evolutionary standpoint, there are still two types of benevolent behavior that evolution can explain: that which we regularly show towards our family, and that which we occasionally show towards strangers.

Regarding the first, apparent altruism towards family members might be explained through the notion of kin selection. It begins with the fact that there is an increased survival rate for organisms that care for their kin. Those that don’t will die out. We’ve thus evolved so that I am instinctively inclined to improve the chances for survival of my family or tribe, and not just my own survival. I am fighting to preserve my genes, and not simply myself. It then may be very natural for me to make major sacrifices for my children who will perpetuate my genes. As to the second, apparent altruism towards strangers might be explained through the concept of soft-core altruism. On this view I am essentially selfish, but I will be kind to other people when it aids my own survival in the long run. This involves figuring out who is on my side, who is against me, and making alliances with the right group at the right time. I may not be fully conscious of exactly why I make the partnerships that I do. One evolutionary biologist argued that even the revered Mother Theresa benefited personally from her association with the Catholic Church, and her seemingly altruistic actions were a part of the arrangement. This is far removed from the genuine altruism that Butler envisioned, but it nevertheless preserves the idea of kindness towards others within the framework of human evolution.

C. Reason and Emotion

Sometimes people don’t want to talk about politics or religion because of the strong emotions that those topics generate. Many moral controversies are also like this, such as abortion, homosexuality, capital punishment, and animal rights. Many animal rights advocates, for example, are so enraged by current practices towards animals that they will vandalize fur stores, break into laboratories and set free animals used in experiments, or picket slaughter houses. The animal rights group PETA—People for the Ethical Treatment of Animals—launched a project called “Holocaust on your Plate.” One of its posters showed Nazi concentration-camp prisoners crammed together in bunks and compared this to chickens jam-packed into cages on a modern factory egg farm. Enraged Jewish groups denounced the project for trivializing the suffering of Jewish victims. Without question there is some connection between morality and emotion. Some
philosophers have gone so far as to say that moral assessments are only expressions of our feelings. On the other hand, other philosophers have staunchly maintained that morality is fundamentally a matter of human reason and not emotion. For them, emotional appeals, such as those by PETA, are completely irrelevant to moral decision making, and we need to approach the subject calmly and rationally.

**Moral Reasoning: Detecting Truth and Motivating Behavior.**

Many philosophers identify two distinct features of moral reasoning: (1) it discovers moral truths, and (2) it motivates us to abide by moral standards. Concerning the first, philosophers have proposed different ways of discovering moral truths. We’ve already looked at Plato’s view that moral truths reside in the spirit-realm, which we access through our reason. For Plato, when I try to access the spirit-realm, I must look beyond the distorted physical world down here on earth and have something like a mystical experience. I mentally reach into the heavens and grasp moral notions, in much the same manner as religious believers look to the heavens for divine guidance. Many philosophers agreed with Plato that moral reasoning was a sort of religious experience, but others downplayed that angle. Instead, they argued, moral reasoning is an ability to recognize moral laws that God implanted in human nature. There is no mystical experience, but only an awareness of instinctive moral principles. Moral intuitions, they maintained, are much like inborn conceptions of mathematics: to grasp them we only need to be attentive to the voice of reason. Whether acquired through a mystical experience or a natural intuition, though, the point is still the same: moral reasoning gives us access to moral truths.

Consider next the second feature of moral reasoning—that it motivates us to abide by moral values. It doesn’t do me much good to discover moral truth and devise a moral game plan if in the end I don’t follow these things. Everything in morality leads to proper behavior, and without the right motivation, my best reasoning on an issue remains only theoretical. Many philosophers argue that moral reasoning in itself motivates us to do the right thing. There are in fact several motivations that influence my conduct. Suppose that I am considering lying to my boss about being sick, just so I can get the day off. My motivations for skipping might be laziness and the desire to do something more entertaining. My motivation for not skipping might be fears about getting caught. Moral reasoning, though, puts forward one additional motivation: reason tells us that it is wrong to lie. To the extent that human beings are rational animals, I will be at least somewhat motivated by this rationally-supplied moral principle. Some philosophers felt that moral reasoning is the only motive that matters when making a genuinely moral choice. Emotional motivations, they argue, misguide us from the right course of action.

**The Is-Ought Problem.**

Hume was among the first philosophers to seriously challenge the notion of moral reasoning. He was familiar with standard theories on the subject, but felt that morality was so intertwined with emotion, that there was almost no room left for reason. He targeted both claims above, that human reason discovers eternal moral truths and also motivates us to be moral. He argued on the contrary that human emotion is responsible for both of these tasks. Regarding principle 1—the rational discovery of moral truth—Hume asks you to perform a mental experiment. Think about any horrible deed, such as a murder, and try to discover any special fact about it that constitutes its immorality. After all, the job of reason is to discover facts, and if there was some factual moral truth surrounding murder, surely your reason would spot it. What, though, do you actually find? You will certainly not perceive any eternal moral truth at play. All you will see is an arrangement of bodily movements, feelings, motives and thoughts, but you will never rationally detect the immorality itself. Instead, he argues, “you must turn your reflection into your own breast,” and find an emotion of disapproval within you. Your emotional reaction alone,
then, constitutes the moral assessment regarding the murder.  
Hume’s basic argument is this:

(1) If reason discovered moral truths, then we would be able to  
identify a uniquely immoral factual quality in an action  
such as murder.

(2) We cannot identify any such factual quality, but will only  
find our emotional reaction.

(3) Therefore, reason does not discover moral truths, and,  
instead, all moral assessments are emotional reactions.

Moral assessments, he believes, are much like our evaluations of  
the artistic beauty that we might find in a painting: “it lies in  
yourself, and not in the object.”

Regarding principle 2—that reason motivates us to be moral—  
Hume argues that this position rests on bad psychology. Human  
reason has no ability whatsoever to influence human actions;  
reason simply provides us with facts, but does not tell us what  
we should actually do with them. Imagine that I could prevent a  
nuclear explosion, but in the process of disarming the bomb I  
would get a tiny scratch on my finger. Of course, I should go ahead  
and stop the explosion, regardless of the scratch. However, Hume  
argues, it is our emotions that incline us to make this choice; reason  
doesn’t care one way or the other. I must first desire to place the  
safety of the world above the safety of my finger. So too with all  
moral choices, such as donating to charity. Pile on as many reasons  
for charitable behavior that you like: charity is an intuitive  
obligation; charity follows from the Golden Rule; charity makes  
other people happy. We still won’t be motivated to act charitably  
unless we are emotionally inclined to do so.

Hume’s attack on moral reasoning is encapsulated in the famous  
motto that we cannot derive ought from is. That is, I can’t simply  
conclude that I have a moral obligation (an ought statement) on the  
basis of rational facts that I’m presented with (is statements).  
Suppose that I made the following argument:

Premise: charity is a human instinct.

Conclusion: therefore we ought to be charitable.

My premise is a factual statement about what is the case, and my  
conclusion is a statement of obligation about what we ought to do.  
The problem, as Hume sees it, is that even if charity is a human  
instinct, this does not necessarily mean that we are morally  
obligated to donate to charity. Vengeance is also a human instinct,  
but it is one which we should suppress. Moral obligation comes  
from our emotional reactions, and not from a mere presentation of  
facts.

Moral Utterances Express Feelings. British philosopher A. J.  
Ayer (1910-1989) agreed with Hume, but felt that  
the attack on moral reasoning could be pushed a  
little further. Ayer asks us to distinguish between  
two kinds of utterances, namely, factual reports and  
nonfactual expressions:

Factual reports:
   I’m a fan of that team.
   A cup of rotten milk is on the table.
   I had fun in high school.

Nonfactual expressions:
   Go team go!
   Rotten milk, yuck!
   Oh, for the good old days of high school!
The difference between these two groups is that factual reports are either true or false statements about the world, and nonfactual expressions aren’t. That is, for each of these factual reports we can meaningfully ask whether they are true or false, such as “Is it true or false that I’m a fan of that team?” “Is it true or false that a cup of rotten milk is on the table?” With nonfactual expressions, though, it’s completely nonsense to ask, for example, “Is it true or false go team go?” “Is it true or false rotten milk, yuck?” Nonfactual expressions merely vent our feelings and don’t report any facts at all.

Ayer then asks us to decide in which of these two groups moral utterances should go. Suppose that I say “Donating to charity is a good thing.” At first glance, this appears to be a factual report, since it seems that we can meaningfully ask “Is it true or false that donating to charity is a good thing?” Ayer warns, though, that we should not be fooled by first appearances. Moral utterances like this really belong in the second category of nonfactual expressions. When I say “Donating to charity is a good thing,” what I really mean is something like “Hooray for Charity!” Similarly, when I say “Stealing is a bad thing,” I really mean “Boo for stealing!” Moral utterances merely vent feelings, they report nothing factual at all. Ayer’s position is called emotivism: moral utterances express feelings, but do not report facts.

Emotivism attacks moral reasoning at an entirely new level. Philosophers in the past commonly felt that utterances such as “Donate to charity is a good thing” are factual reports that we make through the use of our reason. For Plato, this utterance reports that “charity is an eternal moral truth.” For Sextus Empiricus, it reports that “society approves of charity.” Even Hume might take this to mean “I approve of charity”—which is a report about one’s personal feelings. For Ayer, though, moral utterances don’t even rise to the level of factual reports about our feelings. They are emotional hisses, boos, hoorays and bravos, much like an animal might express when excited.

Many philosophers today are reluctant to embrace Hume’s and Ayer’s dismal assessment of moral reasoning. No doubt, there are genuine limits to human reasoning, but in moral matters it seems that reason does provide some kind of guide. Even if there is some emotional component to moral assessments, we still rely on reason to sift through the facts and weigh the arguments pro and contra on various controversies. For example, are chicken coops really like Nazi concentration camps as PETA would have us believe? Reason certainly has something to say on this matter. On the one hand, there is a similarity to the degree that in both cases conscious living creatures are painfully crammed into close quarters against their preferences. On the other hand, though, there is an enormous difference between the mental capacities of chickens and humans, and concentration camp prisoners experience pain on many more levels than do pent up chickens. So, while it may be wrong to inflict any pain on chickens, the comparison with concentration camp prisoners is greatly exaggerated. And that’s what reason tells us on this issue without the aid of emotion.

D. Virtues

Philosophers of the ancient world were consumed with the idea of developing moral character. We read in the book of Proverbs that “The wise person conceals his knowledge, but the foolish person blurts out nonsense.” This recommends that we cultivate the character trait of wisdom and flee from foolishness. Confucius notes four characteristics of a superior person: “when conducting himself he is humble; when serving superiors he is respectful; when helping others he is kind; when ordering people he is just.” Here, humility, respect, kindness and justice are qualities that we should adopt. Let’s return to our opening example of the Los Angeles bank robbers and examine it from the perspective of character traits. The
bad guys were greedy in their desire for money, unjust in stealing it from others, and malicious by shooting everyone in sight. The good
guys, on the other hand, exhibited bravery as they were
overpowered by the robbers’ machine-guns. They showed
ingenuity by borrowing more powerful weapons from the local gun
store. They were persistent by battling it out for 45 minutes, not
letting the robbers escape. These good qualities are what we call
virtues and the bad qualities vices. The most influential theory of
virtues was developed by the Greek philosopher Aristotle (384–
322 BCE), which we will examine here.

The Virtuous Mean. Aristotle defined human beings as “rational
animals.” That is, humans have the major biological features that
we also find in dogs and cats, but at the same time we have an extra
psychological ability to reason, which elevates us above the other
animals. Morality, according to Aristotle, is an interesting interplay
between the animalistic and rational components of human nature.
Let’s begin with the animalistic elements. Fido the dog has many
animalistic urges that get him through the day. He gets frightened
when facing danger, he enjoys eating food, he gets angry at
intruders, and he desires companionship. Humans have these same
animalistic drives, and several more that are unique to us. We have
desires to donate money to the needy, to feel good about ourselves,
and to make people laugh. Today we might say that all of these
drives were naturally implanted in both animals and people to aid
in survival. If Fido didn’t get frightened when facing danger, he’d
walk into a bear cave, wander off a cliff, or something that would
quickly put an end to his life. Human drives serve largely the same
function. Even our human desire to amuse others is important for
social acceptance, which in turn aids in our survival.

Much of our human psychology, then, is rooted in purely
animalistic urges. The difference between Fido and me, though, is
that I have the rational ability to control my animalistic drives. I
can train myself to suppress my urges when needed, and act out on
them when the situation demands it. Take, for example, the natural
inclination to get angry. We need this drive to keep people from
harming us. If I never got angry, Aristotle argues, people would
steal my belongings, hit me for amusement, and harass my family.
On the other hand, it’s not good for me to have a short fuse and fly
into a rage if my waiter puts too much ice in my drink. I need to
display the right amount of anger in the right situation.

As a rational creature, then, my job is to develop the right kind of
mental habits that regulate my natural urges. These good mental
habits are moral virtues, and they will be at a happy medium
between two bad habits, or vices. For example, if I properly restrain
my urge to get angry, then I will have the moral virtue of good
temper. If I don’t restrain myself enough, I will have the vice of ill-
temper. If I restrain myself too much and never get angry, I will
have the vice of spiritlessness. Aristotle lists about a dozen virtues
that follow this same formula: virtues that stand at a mean between
vices of deficiency and vices of excess. Here is a chart of some
especially interesting ones:

<table>
<thead>
<tr>
<th>Natural urge</th>
<th>Vice of Deficiency</th>
<th>Virtuous Mean</th>
<th>Vice of Excess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>Spiritlessness</td>
<td>Good temper</td>
<td>Ill-temper</td>
</tr>
<tr>
<td>Fear of danger</td>
<td>Cowardice</td>
<td>Courage</td>
<td>Rashness</td>
</tr>
<tr>
<td>Pleasure</td>
<td>Insensibility</td>
<td>Temperance</td>
<td>Intemperance</td>
</tr>
<tr>
<td>Aiding others</td>
<td>Stinginess</td>
<td>Generosity</td>
<td>Extravagance</td>
</tr>
<tr>
<td>Self-worth</td>
<td>Self-loathing</td>
<td>Self-respect</td>
<td>Arrogance</td>
</tr>
</tbody>
</table>

On the surface, it doesn’t seem too difficult to find the proper
middle ground with all of these virtues. To be courageous, for
example, we just avoid being too cowardly or too rash. However, Aristotle argues that finding that middle ground is easier said than done. Suppose I am a police officer and I understand that courage involves knowing how to regulate my fear of danger. Am I cowardly if I don’t chase down an armed bank robber? And how many bullets does the robber have to fire before my pursuit of him becomes rash? Developing virtues takes much time and training, and often begins in childhood education.

There’s something unusual about Aristotle’s theory: while he tells us everything we need to know about virtues, he says almost nothing about moral rules, such as don’t kill, and don’t steal. He certainly was aware of the function that moral rules play in people’s lives, and he himself had to regulate his conduct by the rules of law within his native city of Athens. He may have simply felt that virtues were the most central part of moral philosophy and thus focused only on that aspect. In any event, in contemporary times it has led to the theory that virtues alone should be the focus of morality, not rules. Let’s call this the virtue-alone theory.

Virtues and Gender. Many feminist philosophers have embraced the virtue-alone view of morality. To see why, consider the age old dispute about whether men and women differ in important psychological ways. We are all familiar with stereotypical male/female differences. Young boys play with toy guns and swords, and roughhouse with each other. Young girls play with doll houses and interact more gently. Stereotypes go beyond aggression, though. When older, men are better at mathematics and other analytical tasks, while women have stronger communication skills. Psychologists dispute about how much of this difference is inborn, and how much is the result of centuries of social conditioning in male-dominated societies. The fact remains, though, that there are at least some psychological marks of distinction between men and women, and some of these may be relevant to theories of morality. Feminist philosophers have noted one major area of distinction: men seem to be fond of categorizing things and inventing rules. Women are less inclined to do this and seem to be more sensitive to the uniqueness of particular situations. Sometimes rule-following may be necessary, as when, for example, designing bridges and skyscrapers. Other times, though, obsession with rules does more harm than good, and this may be the case with rule-based approaches to morality.

The virtue-alone theory addresses the concerns of feminist philosophers in three ways. First, virtue theory is less male-oriented since it downplays rules. Second, the development of virtuous habits is an educational task that requires us to be sensitive to the nuances of our surroundings, which is a specialty of female thinking. Third, virtue theory allows us to introduce new ideals into our value system which are more overtly female in character. One almost indisputable fact about female psychology is that women tend to be more nurturing than men. Women now dominate the fields of counseling, social work, primary education, and nursing. Even with all the changes in modern society, women are still more involved in child-rearing than men. American philosopher Carol Gilligan (b. 1936) has argued that, while women have taken care of men, men themselves “tended to assume or devalue that care” both in their theories of human nature and in their daily lives. Thus, an important female virtue that we can add to Aristotle’s list is care.

Virtues and Rules. While the virtue-alone theory has supporters, critics charge that we can’t dismiss moral rules that easily. They make their case with two main arguments. The first involves a paradox: misused virtues can actually become vices. Imagine that a person possessed a string of virtues, such as intelligence, ingenuity, patience, and calmness. “Ah,” we might say, “this person must certainly be a good human being.” However, these are qualities that we find in the most skilled thieves. In fact, the thief will become
more immoral in proportion as he has these qualities. According to critics, this is precisely why the focus of morality should be on unchanging rules, not on virtues. As long as we follow moral rules such as “don’t steal” then we won’t misuse virtues such as intelligence. Our virtues will still be an asset to our moral conduct. For example, when you act on the moral rule “help others in need” you can draw your virtue of intelligence and find the most effective way of assisting the needy. Your primary focus, though, should be on the rule, and not the virtue that assists you in carrying it out.

The second argument against the virtue-alone theory asks us to examine what takes place when we morally praise or blame someone. Suppose, for example, that I get drunk and run over someone with my car. I’m then thrown in jail and the local newspapers make me look like a horrible villain. “But wait,” I say, “I’m actually a very good man. I have all of the virtues, no vices, and in fact I don’t even drink. This was a special occasion and, in the midst of a celebration, I had a lapse of judgment which I never had before.” While my explanation might make me seem less villainous and may even get me some sympathy, I will still be morally blamed for my action and punished. Moral praise and blame, according to the critic, is really a question of whether our actions conform to moral rules, and it is not a question of whether we possess moral virtues. The principal reason is that everything you know about me is based on my actions. You can’t read my mind to see what mental habits are ingrained in my psyche. You only see my actions, and whether they conform to the right rules. Ethics ultimately involves a rigorous science according to which we discover these rules. By comparison, speculating about virtuous habits is child’s play.

What should we conclude about this battle between rule-based and virtue-alone morality? First, an exclusive focus on moral rules does seem to distort the age-old tradition of moral thinking, and some philosophers seem to have erred by eliminating virtues from their moral theories. However, we may rightfully question whether moral rules should be entirely replaced with virtues, as some virtue theorists have suggested. Just as virtues are a longstanding part of our moral tradition, so too are moral rules, and for thousands of years the two approaches have harmoniously coexisted. Second, it seems clear that Aristotle’s original list of virtues needs amending. Aristotle felt that women and slaves held a lower place in the human social order and, so, his list of moral virtues has the feel of an aristocratic male. The list of virtues that we endorse today should certainly reflect our growing traditions of human equality. To the extent that nurturing has been an undervalued female characteristic, it makes sense to acknowledge the virtue of care. However, we may also rightfully question whether moral theories should be predominantly female in character, with its sole emphasis on care, as some writers have suggested. Male/female differences aside, human beings are multifaceted and we may never find a single virtuous character trait that fully encompasses the complex nature of moral obligation.

### E. Duties

Just as we can examine the morality of the Los Angeles bank robbery in terms of virtues and vices, we can similarly list the moral rules that the robbers violated. The two obvious ones are “Don’t kill” and “Don’t steal.” These are obligations—or duties—that are acknowledged in cultures throughout the world. A 4,000 year old Babylonian text succinctly lists our principal moral duties: “Has he intruded upon his neighbor’s house, approached his neighbor’s wife, shed his neighbor’s blood, stolen his neighbor’s garment?” We find these basic principles in the Ten Commandments and in other religious codes of ethics. We don’t need to prove the authority of these duties; we somehow naturally accept them as facts. This is the basis of a theory of morality called **duty theory**.
Duties to God, Oneself and Others. The basic elements of modern duty theory were developed by German philosopher Samuel Pufendorf (1632-1694). God, he argues, has implanted a natural sense of morality within us all, which gives us the precise guidance we need for proper social interaction. These instinctive moral principles are as natural to us as the ability to speak languages, and are so firmly rooted in our minds that we can’t wipe them out. We discover and understand them through the use of our reason, and, to that extent, morality is embedded in the rational part of human nature. All of our moral obligations, according to Pufendorf, are of three types: duties to God, oneself, and others. Let’s briefly look at each of these groups.

According to Pufendorf, we have two principal duties to God: know that God exists and obey God. Like many philosophers of his time, Pufendorf believed that God’s existence could be rationally proven—such as by observing complex design within nature and inferring that it must have been produced by a cosmic designer. Pufendorf feels that this kind of knowledge of God is not optional for us, but is actually a moral requirement. Next, we have a moral obligation to obey God in various ways, such as honoring him, worshiping him, and praying to him. While religious thinkers of the time agreed wholeheartedly with Pufendorf’s view of duties to God, in later centuries most philosophers were more cautious on this subject. Even if God exists and we have some obligation to him, it seems to be more of a religious duty than a moral one.

Pufendorf believed that human nature has both a mental and a physical component, and, so, our duties to ourselves fall into these two groups. Regarding our duties to our minds, he argues that we should develop our talents. Learn a trade, write a book, play an instrument; it’s up to us to decide what we should do, but we should choose something. Even if I can live comfortably off inheritance money without ever working, I am still under a moral obligation to develop some ability. Regarding duties to our bodies, we shouldn’t intentionally do things that cause us physical harm and, most seriously, we should not commit suicide under any circumstance. Like duties to God, philosophers today are a little suspicious about the existence of moral duties to oneself. Don’t I have a right to be a lazy couch potato as long as I’m not a burden to others? Don’t I have a right to die if I’m terminally ill and in intense pain?

The most important—and least controversial—group of moral duties is the final one, namely, duties to others. Here we find the usual obligations, such as don’t steal, murder, or lie. These also fall into different subgroups. There are, for example, special duties that we have to our families, such as caring for our children, being faithful to our spouses, and respecting our parents. We have special duties to the people in our local communities, such as being charitable to them, and other obligations to our governments, such as obeying the laws.

The Categorical Imperative. Pufendorf’s duty theory was ultimately eclipsed by that of German philosopher Immanuel Kant (1724–1804). What remains important about Pufendorf’s theory, though, is the idea that morality involves instinctive moral obligations that are unwavering and do not depend upon our private desires. Kant believed that Pufendorf’s theory was on the whole correct, but argued that we don’t need to memorize a long list of distinct duties to God, oneself and others. At bottom, Kant maintained, there is only a single principle of moral obligation, which he dubbed the categorical imperative. All of our distinct duties, as valid as they may be, are just applications of this.

Before telling us exactly what the categorical imperative is, Kant explains how moral judgments differ from other value judgments that are less morally crucial. Throughout the day people tell us that
we ought to do certain things. We ought to read some book, we ought to take better care of our health, we ought to donate to some charity. Although all of these injunctions may involve some kind of obligation upon us, few of these are genuinely moral obligations. How can we tell the difference? According to Kant, non-moral obligations are merely hypothetical imperatives: they tell us that if we want to achieve some goal, then we should perform some act. If I want to lose weight, then I should go on a diet. If I want to be entertained, then I should read a specific book. The reason that these are non-moral obligations is because nothing requires me to have those specific goals. Maybe I don’t want to lose weight. Maybe I’m not interested in literary entertainment. Moral obligations are entirely different, though. They do not depend on our personal preferences and are obligatory no matter what. Truly moral obligations are categorical imperatives—that is, absolute commands. “Don’t kill,” “Don’t murder,” and “Don’t lie” are certainly of this sort.

The ultimate categorical imperative, for Kant, is a very broad moral principle that we can apply in every circumstance. He paradoxically gives four versions of it, claiming that they say the same thing from different perspectives.

[All of the versions are based on the assumption that we should focus on motive or intention in evaluating the morality of an action. Kant’s reason: we cannot be obligated to act in a certain way unless everyone else is likewise obligated. The fact that no one is exempt from this obligation is what makes an action a moral duty. That is why Kant calls his ethical imperative categorical (i.e., universal, admitting of no exceptions). Being moral means, then, acting in a way by means of which we intend or will to be guided by rules that everyone can and should adopt.]

One version of Kant’s categorical imperative is especially powerful: treat people as an end and never only as a means to an end. Most simply, this means that we should treat people with dignity, and not as mere objects. Behind this principle is a crucial distinction between two kinds of value that we find in things: instrumental and intrinsic. Something has instrumental value when it is a tool to accomplish something else. My house key, for example, is valuable since it opens my front door. If I change the lock, the key no longer serves a purpose and I toss it in the garbage. Most things that we see have only instrumental value. Our jobs are valuable because they give us money. Money is valuable since it allows us to buy stuff. By contrast, intrinsically-valuable things are important in and of themselves, regardless of what further benefit they have. Happiness, for example, is intrinsically valuable. If I say “I’m happy right now,” it would be odd for you to ask “what further benefit do you get from being happy?” Happiness is simply valuable for its own sake. Armed with this distinction between instrumental and intrinsic value, Kant’s categorical imperative tells us that we should not use people as instruments for our own benefit, like house keys, but treat them as intrinsically valuable beings, just as we would view our own happiness. People are intrinsically valuable, Kant thinks, because we have free wills and can shape our surroundings according to our own design. We can’t say that about a house key. Because we are unique in this respect, we must treat people with a special dignity.

The beauty of this principle is that it tells us very clearly why specific actions are right or wrong. If I donate to charity, I am recognizing the intrinsic value of the person I am donating to. When I develop my talents, I recognize my own intrinsic value. On the other hand, if I steal a pack of gum from the corner store, I am treating the store owner as a mere object for my own benefit. If a man cheats on his wife, he treats both his wife and mistress as mere objects of gratification. Kant envisioned the categorical imperative as a more precise version of the Golden Rule, and, to a large extent the principle achieves that aim.
**Duties to Animals and the Environment.** Ethicists of the past felt that moral obligations applied principally to human beings, and perhaps also to God. Pufendorf, for example, held that God implanted moral obligations in people as a means of enabling us to live in peaceful human societies. Animals were irrelevant to that mission. Kant felt that we only have moral obligations to creatures that are capable of making free and rational decisions. Animals, he believed, can’t do that: they operate only on instinct, and don’t have the mental capacity for anything like rational thought. Not only do animals lack rationality, Kant says, they aren’t even conscious of themselves. Even though they might give the appearance of being in pain, for example, they don’t have a conscious experience of pain itself. While we have no direct duties toward animals themselves, Kant argued, we nevertheless have indirect duties based on how cruelty to animals might impact human beings. Suppose that you routinely set dogs’ tails on fire just for the fun of it. Eventually, that sadistic thrill will wear off and you’ll be inclined to work your way up the food chain and torture humans for your entertainment. Thus, according to Kant, while setting dogs’ tails on fire won’t violate a direct duty to dogs themselves, we have an indirect duty to avoid such conduct because it inclines us to be cruel to humans.

Today this reasoning seems quite bizarre. Even though some animals such as chickens lack the mental capacity for rational thought, they still are conscious of pain, and it is morally wrong to torture them. This suggests that at least some moral duties go a step beyond human beings and apply to any conscious creature with a capacity for experiencing pain. Part of the reason for this shift in attitude since Kant’s day is that we know more about animal physiology than we used to, and we know that even animals like chickens are biologically closer to us than we previously thought. When we move up the animal hierarchy, many creatures—such as dogs, cats and chimpanzees—have even greater mental capacities. They obviously can’t play chess or write computer programs, but some have the rational abilities to solve problems, make tools, and learn languages. Further, they are not just aware of their surroundings, as chickens are, but are also self-aware. That is, they are aware of themselves moving through time and have something like hopes and dreams for the future. Along with these greater mental capacities, then, comes greater moral duties on our part, specifically to respect and to some extent facilitate their goals.

Some of these moral duties towards animals have worked into legal codes—laws about cruel and inhumane treatment of pets, and even regulations about how animals should be treated in slaughterhouses. The question today really isn’t so much of whether we have moral responsibilities towards animals, but how far those duties extend. Is it wrong to eat or experiment on animals, just as it is wrong to do with humans? These issues are still up for debate.

Pushing the issue of non-human duties even further, we can ask whether we have any obligations to plants as well as animals. Plants clearly aren’t conscious, and it is impossible to inflict them with pain as we might a chicken. But we have growing concerns about the depletion of rain forests and the mass extinction of thousands of plant species on a regular basis. There may be nothing wrong with killing a bunch of weeds in my back yard, but if we slash and burn the last members of an endangered plant species, that’s a different story. As with duties towards animals, though, there are two ways that we can look at our responsibility towards plant species.

First, maybe our duties to the environment are only indirect, based on how human beings will be adversely affected by environmental irresponsibility. If we chop down all of the rainforests, the earth’s temperature will skyrocket, and we’ll all die. If we wipe out too many plant species, we disrupt the food chain, and we starve. Even if we are not particularly fond of plants, there are lots of human-centered reasons to acknowledge an indirect duty towards them. Second, maybe our duties to the environment are direct, and we have moral responsibilities to plant species themselves for their
own sake, irrespective of their impact on human interests. On this view, plant species in and of themselves have a moral standing, similar to the way that humans do, and this generates a direct moral duty towards them by us. Direct duties to the environment, though, is a tough sell. It doesn’t seem like an obvious moral duty, on the same level as duties to humans or even duties to higher animals. Advocates of this ecological approach often draw attention to the interconnectedness of all things, and point out that it doesn’t make sense to talk about human duties and ignore the surrounding fabric of life into which humans are woven. Some advocates believe that it requires an environmental awakening—a kind of mystical experience—for a person to grasp the independent moral standing of plant species and environmental systems. But people who are incapable of having that kind of experience may be left scratching their heads in wonder.

**F. Utilitarianism**

Duty theory seems like a good way of understanding our moral responsibilities—to people, animals and the environment. What could be wrong with it? The problem, as critics see it, is that these so-called instinctive duties are pure fabrications. Pufendorf felt that God permanently implanted duties in our nature; Kant felt that duties are an integral part of human reasoning. They are not, though, as fixed and universal as these duty theorists have maintained, which we see most evidently in the various duties that have come and gone over the centuries. Once considered essential to morality, duties to God have been cast aside. Suicide used to be among the top moral crimes, but now we have a right to die. Now we have newly-discovered duties to animals and the environment. The concept of instinctive duties seems to be just a sophisticated justification for personal convictions, or, worse yet, for personal prejudices. Wouldn’t it be great if we could arrive at our moral principles more objectively? More scientifically?

In the late 18th century, a philosophical movement called *utilitarianism* aimed at doing just that. The concept was simple: we measure right and wrong by considering the pleasing and painful consequences of our behavior upon ourselves and others. We no longer have to root through human instincts to discover our alleged duties. Instead, we simply examine the publicly observable consequences of our actions, and tally the pleasures and pains that result. What the Los Angeles bank robbers did was wrong, for example, because their actions produced far more pain than pleasure. They themselves died painfully, and caused extreme pain to those who they shot. By contrast, the charitable acts that Mother Theresa did throughout her life were morally good because they produced great pleasures for the needy, which overbalanced the modest pains that she herself endured through self-sacrifice.

*The Utilitarian Calculus.* British philosopher Jeremy Bentham (1748-1832) was an important developer of this theory. A stickler for details, he argued that it wasn’t good enough to simply have a gut feeling about whether an action produced more pleasure than pain. To be scientific, we must *quantify*—that is, assign number values to—the pleasing and painful consequences of an action. We then tally the numbers, and assess the total score. Suppose that the Los Angeles bank robbers produced 100 total units of pain, and only 5 units of pleasure—pleasure, perhaps, from the money that gun store owners received from the sale of the weapons and ammunition. Since the “pain” score is higher than the “pleasure” score, we thereby pronounce the bank robbers’ conduct as morally wrong. Bentham called this approach the *utilitarian calculus.*

Not all pleasures and pains carry the same weight, he argues, and to be accurate we must account for their differences. For example, the pleasure I receive from eating a slice of pizza is not as weighty as the pleasure of winning a million dollar lottery. Bentham noted
seven important factors in assessing the differences between various pleasures and pains:

(1) Intensity: how extreme the pleasures and pains are.
(2) Duration: how long the pleasures and pains last.
(3) Certainty: whether the pleasurable and painful consequences are certain or only probable.
(4) Remoteness: whether the pleasures and pains are immediate or in the distant future.
(5) Fruitfulness: whether similar pleasures and pains will follow.
(6) Purity: whether the pleasure is mixed with pain.
(7) Extent: whether other people experience pleasure or pain.

Thus, a pleasure gets more points if it is intense, long, certain, and immediate. For example, the pleasures from an amusement park visit would be intense, certain and immediate, but comparatively short. We add more points if that action produces further pleasures down the road, such as the pleasing memories I might have next year of the amusement park visit. We deduct points for pains that are mixed with pleasures, such as nausea from too many rides on the rollercoaster. Finally, for each person that is affected by an action, we repeat steps 1-6. For example, I might get sick on the rollercoaster and throw up on other riders. We must then analyze the intensity, duration, certainty, remoteness, fruitfulness and purity of the pain I inflicted on those riders.

Bentham believed that his utilitarian calculus could have a sweeping impact on society, both with how we determine morality in our daily lives and how government officials would craft legislation for the betterment of everyone. Gone would be the days when are actions and social policies would be guided by mere moral hunches. The utilitarian calculus gives us a precise formula for making moral decisions based on hard facts.

**Higher Pleasures and Rules.** Bentham was the godfather and teacher of a child prodigy, John Stuart Mill (1806–1873). As a young man, Mill adopted Bentham’s vision of utilitarianism and valiantly defended it against a growing number of criticisms. When a little older and wiser, though, he began finding fault with his teacher’s views, and went so far as to say that Bentham was not a particularly good moral theorist. Mill continued to embrace utilitarianism—and in fact became the most famous 19th century advocate of the theory. However, he departed from Bentham in two important ways.

First, Mill rejected the mathematical approach of Bentham’s utilitarian calculus. In fact, Mill argued, it is impossible to assign numerical values to all pleasures. Some pleasures, Mill conceded, can be quantified. For example, on a scale of 1 to 10, I can assign a value of “2” to the pleasure I get from eating a slice of pizza and a “4” to a visit to the beach. These, though, are merely bodily pleasures that appeal to the lower and baser features of human nature. As rational creatures, though, we are capable of experiencing higher mental pleasures, such as the joys of playing chess, listening to fine music, and viewing great works of art. We can also take pleasure in designing bridges, planning cities and solving social problems like poverty, hunger and illness. These pleasures, though, cannot be measured and plugged into a numerical calculation. We all recognize their merit, and in fact, Mill says, we value them more than bodily pleasures. So, when determining whether something is right or wrong, we survey the consequences and give greater weight to the higher and more dignifying pleasures that result. There are no numbers to tally or scores to compare. It is more of an intuitive assessment.
Second, Mill questioned whether we really need to evaluate the consequences of each one of our actions individually. Bentham’s utilitarian calculus directs us to tally the pleasure and pain of every action that we perform—a position called act utilitarianism. The problem with this is that we don’t have time to morally evaluate each of our actions throughout the day. Life would grind to a halt. Instead, Mill argues, we should just keep following moral rules as we have been doing: don’t kill, don’t steal, don’t lie. However, we should submit those rules for utilitarian evaluation. That is, we must evaluate whether those rules bring about more pleasure than pain.

Consider these two possible rules: (1) Don’t steal, and (2) It’s OK to steal when you feel like it. Assuming that we’d follow them consistently, we should ask which rule would bring about the greatest amount of pleasure in society. Clearly it would be the first. If we allowed people to steal when they felt like it—with no moral or legal penalties—the concept of property ownership would go out the window. You couldn’t open a business because people would just walk in and take what they want. You couldn’t plant food in your back yard for the same reason. For that matter, you couldn’t even count on having a house to live in since every time you’d leave it, you’d be at risk of someone else moving in. On balance, adopting the second rule would have disastrous consequences. Mill’s rule-based approach is representative of a position called rule utilitarianism.

According to Mill, most of our moral rules today have developed through a kind of utilitarian trial and error. Thousands of years ago people adopted the moral rules that clearly benefited society the most. We don’t really have to reinvent the wheel and create our favorite moral principles all over again. In fact, Mill argues, we would rarely ever need to evaluate a moral rule on utilitarian grounds. One exception might be when newer and more specific issues come along. For example, should health care costs be paid for by the government? Should we ban ownership of assault weapons? The answer in each of these cases would rest on which rule would produce the most pleasure. Another exception would be in situations when two moral rules conflict and we can’t follow both. The most famous illustration of this is the case of the inquiring murderer. Suppose that you see a man run down the road, and then jump into a dumpster. Another man carrying a gun comes around the corner and asks, “Did you just see someone run through here?” It’s clear to you that the guy in the dumpster is a dead man if you tell the truth. You are now caught in a dilemma between two moral principles: don’t lie, and don’t cause harm to others. According to Mill, you should resolve this specific conflict as an act utilitarian would: determine which course of action (as a rule) would produce the most pleasure and least pain. Clearly, the best course of action here is to lie to the would-be killer and avoid causing harm to the man in the dumpster. Once the situation is resolved, you return to following the usual moral rules.

Reactions from Duty Theorists. From the time that utilitarianism first emerged, duty theorists were unimpressed with its so-called scientific approach. Two specific problems were raised, which even today remain among the most serious challenges to utilitarianism. First, there’s no question that utilitarianism aims at a precision far beyond what duty theorists ever dreamed of. The problem, though, is that its standard of precision is too high. Could we ever fully evaluate every positive and negative consequence of the Los Angeles bank robbers’ conduct? Once the story hit the media, millions of people had reactions to it, which we could never fully calculate. Many of these reactions were no doubt painful feelings of revulsion towards the robbers’ brutal conduct. But these feelings would have been mixed with pleasing feelings from the entertaining nature of the story itself. A few years later the story was made into a movie, which, undoubtedly made money for some people and thus produced even more pleasure. For all we know, two actors on the movie set fell in love and had a child who will grow up and
some day discover a cure for cancer. We just don’t know, and we are not in a position to precisely calculate all the consequences. One early critic of utilitarianism held that even the wisest of people will only ever have a faint glimpse of the consequences of an action. “The nature of general consequences,” he argued, “is too comprehensive to be embraced by the human understanding, too dark to be penetrated by human discernment.” Rule utilitarianism attempts to sidestep this issue by having us look at rules, such as “Don’t steal” rather than specific actions such as those of the Los Angeles bank robbers. Thus, we don’t need to predict every consequence of every action. The problem re-emerges, though, when we evaluate the consequences of rules that we might want to adopt. The long-range consequences of any moral rule are well beyond our ability to grasp.

Second, utilitarian reasoning may lead us to adopt actions or rules that conflict with important traditional values. Suppose that I kidnap you and make you my slave. I’ll have you mow the yard, clean out the cat box, overhaul the engine in my car, and any other unpleasant task that I can think of. I’ll also share you with my neighbors so that you can relieve them of their unpleasant tasks as well. Your life will certainly be miserable, but through your services our lives will be considerably happier and might actually outweigh your misery. So, on utilitarian reasoning, enslaving you is the morally right thing to do. Again, rule utilitarians attempt to address this problem by focusing on rules rather than actions. They argue that we need to consider the consequences of adopting rules regarding slavery, and a rule allowing slavery would produce more pain than pleasure. However, critics argue, a carefully crafted rule permitting slavery might produce more pleasure than pain. What if we enslaved only people with docile personalities and then sterilized them before they could reproduce? They would be more content in their condition than the average person would be, and we would prevent slavery from becoming hereditary. Duty theorists would object that this kind of slavery would still be wrong, regardless of the pleasure/pain tally.

As we’ve seen, the key motivation for adopting utilitarianism is that it frees us from unreliable and prejudicial moral duties that are allegedly grounded in reason and instinct. It places morality squarely in the arena of public observation where we can impartially analyze our actions or rules. This, at least, is what utilitarianism hopes to achieve. Does it accomplish this? Far from it, duty theorists argue. The data that utilitarians need for their evaluation will always be fragmentary and so their conclusions will be flawed. And even utilitarianisms’ best evaluations may take us far from traditional notions of morality. Utilitarians, thus, seem to be getting exactly what they hoped for: they’ve freed themselves from intuitive duties, but in the process have created a frightening system that allows the ends to justify the means.

This is the heart of the debate between duty theory and utilitarianism, and it highlights the difficulty in choosing between any rival group of moral theories. Each of the moral theories that we’ve examined in this chapter has at least some appeal, and it’s difficult to decide which to cast off. What makes matters worse, moral philosophers today are often especially loyal to their favorite theory, and valiantly argue against rival views. But it doesn’t have to be that way, and, historically, it hasn’t always been that way. The major wars between moral theories broke out with Bentham and Kant in the late 18th century, but prior to that, philosophers were happy to adopt a wide variety of approaches and tried to fit them into a larger master plan of moral theory. What mattered was whether a given approach to morality reasonably explained some component of the moral decision-making process. An eclectic attitude like this could also serve us well today in assessing seemingly rival theories. We’ve already looked at compromise approaches to the contest between moral objectivism and relativism. We’ve also seen that contemporary sociobiology
assumes that people are inherently selfish, yet at the same time there human instincts shaped by evolution that have us behave kindly towards others when such arrangements benefit our survival. Finally, we’ve seen that moral virtues can be integrated very will with moral rules: we lay out the basic moral rules of society, and then develop habits to follow those rules spontaneously.

With the battle between duty theory and utilitarianism, there is also room for compromise, and here is one possible approach. Perhaps moral duties are not permanently fixed in human nature as Pufendorf and Kant maintained. Instead, maybe our duties are only social creations that are imprinted on our minds when young, and give us a sense of conviction that lasts throughout our lives. These duties feel permanent and instinctive, but are really not. As society’s preferences change throughout time, our duties occasionally need upgrading, and utilitarian reasoning comes in handy here. For example, we now recognize duties to the environment, partly because we see the negative consequences of environmentally damaging practices. Governments are now taking on the duties of paying health care costs, partly because we see the negative consequences of privately-funded systems. Thus, utilitarianism might serve as a mechanism for reforming the traditional duties that society imprints on us.

For Review

1. Describe Plato’s theory of the forms and how it relates to the moral relativism/objectivism debate.
2. What is Hobbes’s explanation of the selfish motivations behind pity and charity?
3. In what way, according to Butler, is Hobbes’s egoism too simplistic?
4. Explain the expression “we cannot derive ought from is” and how this relates to the controversy between moral reason and emotion.
5. How would Ayer understand the expression “donating to charity is a good thing”?
6. Explain Aristotle’s view of the virtue and vices that are related to our natural inclination to get angry.
7. According to feminists, what is the key distinctive feature of male-oriented theories of morality?
8. According to Kant’s categorical imperative, why is it wrong to steal?
9. According to Kant, why are our duties to animals only indirect?
10. What is Bentham’s utilitarian calculus?
11. According to Mill, why can’t we assign numerical values to some pleasures?
12. Explain the distinction between act and rule Utilitarianism.

For Analysis

1. Try to defend the theory of moral objectivism without relying on a spiritual realm of the forms or the will of God.
2. Evolutionary biologists typically think that genuine altruism to others would not be conducive to an organism’s survival. Defend the notion of genuine altruism against this criticism.
3. Think of a solution to the controversy between reason and emotion that incorporates both sides.
4. One criticism of virtue theory discussed in the chapter is that misused virtues can become vices. Defend virtue theory against this attack.
5. Think of a solution to the controversy between duty theory and utilitarianism, other than the one mentioned at the close of the chapter.
REFERENCES AND FURTHER READING

Works Cited in Order of Appearance


Further Reading


CHAPTER 7: POLITICAL PHILOSOPHY

A. The Social Contract
   Ending the State of Nature
   The Prisoner’s Dilemma
   Social Contracts and Bigotry
B. Rights
   Natural Rights and Revolution
   Are Natural Rights Grounded in Fact?
   Do We Need Rights?
C. Political Liberalism and Property
   Libertarianism
   Welfare Liberalism
D. Individual and Community
   Plato’s Republic
   Communism
E. Governmental Coercion
   Four Justifications
   Liberty and Harm
F. War
   Just War Theory
   Pacifism

For Reflection

1. What would society be like if there were no governments?
2. Do the rights of life, liberty and the pursuit of happiness guarantee us anything concrete?
3. Is the government justified in taxing us to pay for welfare programs for the needy?
4. What is more important: your individual freedom, or your role in the community?
5. Should the government prevent you from harming yourself in a dangerous sport, even if your conduct doesn’t harm anyone else?
6. What sorts of wars are morally justifiable?

An organization was formed some years ago called the “Interim Government of the Republic of Texas” whose aim was to move Texas towards independence. One member wrote, “Imagine a country where the people have no freedom—where a corrupt government controls every aspect of life. How would your life be in a country like that? The government would regulate everything you did.... The people of Texas now have an opportunity for change. We have an opportunity to restore freedom and liberty to Texas.”

The group created a Declaration of Independence, held a Constitutional Convention, elected an interim President, and established a Defense Force. Today they have regular demonstrations and many members have stopped paying taxes. The United States Government was not amused by this experiment in nation-building and responded by arresting participants and confiscating their property.

In spite of efforts like the Republic of Texas, the United States, like many governments around the world, is securely established, and don’t have to think much about justifying their existence. When we do think about the issue, though, some important philosophical questions arise, such as what validates the creation of a government, and what are the limits of its authority? The discipline of political science focuses on the structure of specific governments, such as that of the United States or Canada. Political philosophy, by contrast, looks at the issue more abstractly, focusing on the arguments and principles which justify many governments, not just those of the United States or Canada. In that respect it parallels the philosophy of religion, which looks broadly at the notion of God’s existence, without necessarily referring to any specific religion such as Christianity or Hinduism. In this chapter we will look at the more pressing philosophical issues surrounding governments, such as what justifies their very existence and how much they should meddle in our lives.
A. The Social Contract

In its Declaration of Independence, the Republic of Texas justifies itself in very familiar language:

When a government has ceased to protect the lives, liberty and property of the Texian people, from whom its legitimate powers are derived, ... in such a crisis, the first law of nature, the right of self-preservation ... enjoins it as a right towards themselves, and a sacred obligation to their posterity, to abolish such government, and create another in its stead.

The similarity between this and the United States Declaration of Independence is intentional, and the underlying philosophy of both documents is social contract theory. In a nutshell, this theory states that, to preserve our individual lives, we mutually agree to set aside our hostilities and live in peace under governmental protection. On this view, preserving peace is the only justification for political rule, and, according to some theorists, if a government fails in its assigned task we can abolish it and establish another. There are many versions of social contract theory, but the inspiration for all modern ones is British philosopher Thomas Hobbes (1588-1679).

Ending the State of Nature. Hobbes asks us to imagine life in primitive times before the creation of any governments. We would all be on our own to find food and other necessities, without the protection of law or the police. Two main factors make this a dangerous task. First, life’s necessities are scarce, which creates competition. Suppose that everything I needed in life could be easily acquired by pulling things from a tree: food, clothing, electronic devices would be there in abundance for the picking. No one would be interested in stealing from me since everyone could quickly get identical items by going to the nearest tree. That’s not how it is, though, and as long as we need these things and they’re difficult to come by, we are heading for conflict. Second, we are naturally selfish and thus not inclined to make sacrifices for others in need. Suppose that you and I are hunting for food and come upon a single apple at the same time. If I was naturally generous, I’d be willing to let you have it, or at least agree to split it. Since we’re selfish, though, we’ll both want the whole thing for ourselves and will be prepared to fight for it. This, according to Hobbes, is the state of nature for human beings, and it is nothing less than a war of all against all. I would be suspicious of everyone that I see and act with hostility in an effort to gain a reputation as a tough guy. Traditional notions of morality would be useless, and the most important values I could adopt would be force and deceit. In his most famous passage, Hobbes describes life’s gloomy prospects in the state of nature:

In such condition there is no place for industry, because the fruit thereof is uncertain, and consequently, no culture of the earth, no navigation, nor use of the commodities that may be imported by sea, no commodious building, no instruments of moving and removing such things as require much force, no knowledge of the face of the earth, no account of time, no arts, no letters, no society, and which is worst of all, continual fear and danger of violent death, and the life of man, solitary, poor, nasty, brutish, and short. [Leviathan, 13].

Hobbes’s point is that civilization would grind to a halt as we try to survive. Even planting a vegetable garden would be a futile task since others would just kill me and take what they wanted.

It’s easy to envision a life better than that in the state of nature, and, if possible, all of us would like to have long and prosperous lives. So how do we move beyond the state of nature? Hobbes has a plan which he maps out in three fundamental laws. First, we should seek peace as a means of self-preservation. That is, as rational creatures,
we should recognize that the best way of surviving is to live in peace with others to the extent that this is possible. If we can’t achieve peace, then as a backup plan we can destroy whoever we need to in order to survive, but the more effective plan would be to seek peace. The second law is that we should mutually divest ourselves of hostile rights. In the state of nature, there are no moral restraints whatsoever, and I have complete rights to everything, including my neighbor’s property and even my neighbor’s life. At the same time, though, my neighbor has rights to my property and my life. The key to peace is for me to voluntarily give up my hostile rights towards my neighbor under the condition that he gives up his hostile rights towards me. Thus, I agree not to steal from him under the condition that he agrees to not steal from me. As selfish people, we will be stingy about the rights that we give up, and will only relinquish those that are essential for the peace process. Clearly, giving up hostile rights is essential for peace. This agreement between participants, then, constitutes a “social contract.”

It’s one thing to *agree* to the terms of the social contract, but it’s entirely another to *abide* by those terms. I could, for example, agree to live in peace with you and, once your guard is down, kill you and take your possessions. Law three for Hobbes is that we should indeed keep the agreements that we make. The mechanism for guaranteeing this is the creation of a government. The only way we can make people abide by the contract is to catch violators and punish them. This means establishing a governing body with strong policing powers and the authority to penalize contract breakers. Again, as a selfish person, I will hesitate about relinquishing more personal rights for the creation of such a powerful and potentially threatening government. The sacrifice, though, is worth it when I consider that the alternative is being stuck in the brutal state of nature and dying at an early age.

How powerful does this newly created government need to be? Hobbes argues that it must have absolute authority over citizens in order to keep peace effectively. Any perceived weakness of governmental power will invite contract breakers, which will undermine the peace. Thus, citizens make an unbreakable agreement with one another to invest all authority in the government. Because the government is the expression of the will of its citizens to do all that is necessary for their mutual survival, whatever the government requires of its citizens (short of ordering them to kill themselves) is legitimate. From Hobbes’s perspective, revolutions against a prevailing government are never justified, [because to rebel against the government would be to undermine the very principle by which citizens are free to make choices. So, for Hobbes, the American Revolution a century later would definitely not have been justified.]

**The Prisoner's Dilemma.** In view of the enormous impact that social contract theory has had over the past three hundred years, it is not surprising that it has been a continual target of attack by defenders of rival political theories. As soon as Hobbes’s theory appeared, critics argued that his view of the state of nature was far too pessimistic and, left to their own devices, people would not launch a war of all against all. Maybe an instinctive sense of human kindness would overcome our more selfish inclinations. Maybe human reason would show us the folly of engaging in conflict with others. In any event, critics charged, Hobbes has just not persuasively shown that human nature is as dark as he claims. In the face of such criticism, subsequent defenders of social contract theory painted a kinder, gentler picture of the state of nature, attempting to distance themselves from Hobbes as much as possible.

In recent years, though, many political philosophers have felt that Hobbes actually got the story more or less correct. While in his own time Hobbes could only speculate about how people would behave...
in the state of nature, today we have sophisticated ways investigating these kinds of questions. One approach is based on a tactic that police investigators commonly use with great success, popularly known as the prisoner’s dilemma. Suppose that I and an accomplice named Joe are caught robbing a convenient store. The police investigator has some evidence against us, but it’s too weak for a slam-dunk case. What the investigator needs is a confession from one or both of us. Joe and I are not naive enough to volunteer information, so, applying psychological pressure, the investigator offers a plea bargain depending on who cracks first. The conditions of the plea are these:

- If I confess and Joe does not, then I will get only a 1 year sentence, but Joe will get a 10 year sentence.
- If Joe confesses and I do not, then Joe will only get a 1 year sentence, but I will get a 10 year sentence.
- If neither of us confesses, then we will each get a 2 year sentence.
- If both of us confess, then we will each get a 5 year sentence.

Clearly, the best mutual arrangement would be for Joe and me to keep our mouths closed and just serve two years each. Unfortunately, the investigator has placed Joe and me in separate rooms, and we can’t discuss our options. I’m now wondering what Joe is going to do and he is wondering what I’m going to do. I’m thus on my own and have to decide what the best deal is for me, regardless of how Joe responds. As a rational person motivated by self interest, I see that it is best for me to confess, which will guarantee that I won’t serve the big 10 year sentence.

How does this apply to social contract theory? Suppose that you and I are in the state of nature, standing in front of a single apple. As we wonder which of us will get that apple, we are in much the same situation that Joe and I were in. While you and I might talk about our options for splitting the apple, I don’t really know what is going on inside your mind, and I can’t trust anything you say. So, while splitting the apple might be the best mutual arrangement for us, I have to consider what is best for me irrespective of what you do. I then see that it is in my best interests to attack you and take the apple before you attack me. If you and I are family members, I may very well have a natural sense of kindness towards you, and I may just give you the apple outright. If you are a stranger, though, all bets are off and I’ll always be postured for war, just as Hobbes maintains. So, in the absence of government, life in the state of nature would very likely be “nasty, brutish and short,” as Hobbes words it.

Social Contracts and Bigotry. A recent line of criticism against social contract theory is that the social contract is an exclusive club whose membership is not necessarily open to everyone. Hobbes depicts the typical social contract club member very generically, as a rational person who cuts the best deal with others to obtain peace. In point of fact, critics argue, there is an opportunity for a dominant group to take charge and craft the contract in ways that will be to their own advantage. Men, for example, might negotiate contracts on behalf of themselves, and demote the interests of women. One race might call the shots of another. Animal rights advocates charge that club members might direct contracts for human benefit, disregarding the interests of animals. The excluded groups, then, could be assigned the status of second class citizens or declared the mere personal property of club members; they might even be banished completely from the club, with no political status except that of a potential enemy to be destroyed at will.

Not only is this a problem for the social contract as a political theory, but, critics charge, it is a problem that we find in the actual contractual policies that governments have made throughout history. Governments run by men have routinely created laws that
restrict the economic and social freedoms of women. In colonial times European countries declared the natives of foreign lands to be less human, and thus subject to colonial rule. Laws of countless governments have permitted the torturing and killing of higher animals for the benefit of human labor, nourishment or sport. In the real world, then, the social contract club has been an organization run by oppressive bigots.

The key philosophical question here is whether there is something about an original social contract which allows for the potential domination of bigots. We could of course imagine a contractual situation in which men and women of all races—and spokespeople for higher animals—all had an equal voice in determining the laws of the land. But the chaotic atmosphere of the state of nature itself prevents us from knowing who the first negotiators will be. Dominant groups may very well highjack the negotiation process and leave weaker groups with no choice: join our club in a second-class status, or continue at war with us. Some peace is better than no peace, and, so, a rationally self-interested minority may willingly accept the offer. It seems that Hobbes’s version of social contract theory does not guarantee a society of equals. Fortunately, there are other political theories—and even modified versions of social contract theory—that aim to protect the idea of equality. One of these is rights theory.

B. Rights

Perhaps the most fundamental concept in political thought today is that of rights, and we see people assert rights to just about everything. For example, Hooters, the infamous restaurant chain, even claimed that it had a unique right to use scantily clad women to sell food and beer, and it sued a rival eatery for copying their approach. Most typically, we assert our rights to privacy, speech, and religion, and we claim that our rights are violated when the government or some person obstructs our freedoms. A “right” is best understood as a claim against another person—for example, Hooters’ claim against a rival’s copycat tactics. Sometimes my rights are claims to be free from harm that others might inflict on me, such as being robbed, beaten up, or murdered. Other times my rights are claims to be free to act how I please, such as to speak, write, or travel as I wish. In either case I am staking out a territory of freedom and telling others to leave me alone.

Some of our rights are clearly created by governments: they are voted on by government bodies and can be altered or retracted by the same government. The right to drive at age 16 is a good example: different governments can set different age requirements for driving, and we can’t say that one convention is necessarily morally superior to another. Because of their dependence on legislative processes, these rights are commonly called legal rights. Other rights, though, seem to be independent of governments, and these are commonly called natural rights. This special set of rights has three key features. First, they are natural in the sense that we are born with them. Second, they are universal since all humans world wide possess them. Third, they are equal in that every person regardless of race or gender has them to the same degree.

Natural Rights and Revolution. In spite of the universal nature of moral rights, ironically, the concept of “rights” was invented only about 300 years ago. Our modern view of the subject was forged by British philosopher John Locke (1632–1704). Locke was familiar with Hobbes’s view of the social contract and he uses that theory as a way of introducing the concept of rights. Locke asks, what would things really be like in a state of nature? Would it be moral chaos as Hobbes maintained? No, Locke answers. Even in the state of nature all people have fundamental and God-given rights to life, health, liberty and possessions. That is, I have a natural right not to be killed or physically harmed, to behave as I please, and to acquire property
without people taking it from me. I retain all of these rights, Locke argues, as long as I respect the rights of others. Suppose, though, that I violate your property rights by stealing your lawnmower. According to Locke, I thereby forfeit my own rights and you’re entitled to [punish me for my transgressions. Such punishment must fit the crime and be in keeping with what is just according to the law of nature. However, because those who punish transgressors are often driven by vengeance rather than justice, we need to establish a government to restrain excessive punishments.] The point is, even within the state of nature there are moral constraints on our behavior which are determined by our natural rights.

While the state of nature may not be as ruthless as Hobbes contended, it is still a potentially warring environment since people would still prey on each other, especially if they thought they could get away with it. So, Locke argues, to improve our safety we create governments whose sole responsibility is to protect our natural rights. More precisely, we create an agreement with the government: we give the ruler power and authority over us, in exchange for which the ruler protects our natural rights. Suppose, though, that our government fails to keep its part of the bargain; its police force, for example, may not be up to the task of dealing with organized crime, or the government itself may be so corrupt that it routinely violates our natural rights by unjustly confiscating our property or executing us. Would we be entitled to revolt against the government? Hobbes, we’ve seen, felt that the governments that we create need to have absolute and uncompromising authority over us in order to adequately do their jobs and, so, revolutions would never be justified. Locke disagrees. The deal that we cut with governments does not give them absolute authority; we certainly need to give them enough power for them to do their job, but, a deal is a deal, and if our government fails in its task, then we can remove it and create another.

In the 18th century, Locke’s theory was enormously popular among the British who felt that it offered a perfect justification for the Glorious Revolution of 1688 which deposed their king. In time, Locke’s theory became the philosophical inspiration for other revolutions, most notably the American Revolution of 1776. Locke’s influence is especially evident in the Declaration of Independence, penned by Thomas Jefferson:

We hold these Truths to be self-evident, that all Men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty, and the Pursuit of Happiness—that to secure these Rights, Governments are instituted among Men, deriving their just Powers from the Consent of the Governed, that whenever any Form of Government becomes destructive of these Ends, it is the Right of the People to alter or to abolish it, and to institute new Government, laying its Foundation on such Principles, and organizing its Powers in such Form, as to them shall seem most likely to effect their Safety and Happiness.

Like Locke, Jefferson states that people have God-given natural rights, and governments are created to protect these rights; when a government fails in its assigned job, it can be overthrown and replaced.

Are Natural Rights Grounded in Fact? The success of the American Revolution inspired other revolts, which used a similar justification, namely, protection of natural rights. Although the idea of natural rights is now permanently etched in the political imaginations of people throughout the world, it nevertheless faces conceptual challenges. One problem, voiced by British utilitarian philosopher Jeremy Bentham (1748-1832), is that the concept of natural rights has no foundation in fact, and is only an imaginary fabrication. While legal rights are
completely legitimate, Bentham argues, natural rights are not. In general, “rights” just don’t emerge from thin air, and instead come at the end of a chain with three links: legislators create laws, and these laws grant us rights. For example, governmental legislators create laws allowing 16 year olds to drive, and from these laws qualified 16 year olds have the right to drive. All three of these links—legislators, laws, rights—involv e facts. We can read the legal statutes that define our legal rights; we can visit our Capital and ask the legislators themselves about the laws they enacted. When we turn to natural rights, we find a similar three link chain: a divine legislator creates natural laws, which in turn give rise to natural rights. The problem, though, is that none of the links in this chain are subject to factual scrutiny. There is no official codebook of natural law for me to consult; if I seek clarification from the divine legislator himself, I get silence. While many philosophers have written books in defense of divinely inspired natural law, there is no consensus about what these laws are, and no clear procedure for even investigating the matter. Thus, the concept of natural rights is founded less on fact than it is on imagination and perhaps wishful thinking.

Bentham recommended that we abandon the idea of natural rights and instead limit our notion of rights to the legal arena. We thus can rely on governments to confer on us the rights to life, liberty, happiness, and anything else that it deems appropriate for a happy society. Many political philosophers agree with Bentham, and recent political documents are more cautious about their use of the word “natural.” The most important example of this is the United Nations’ 1948 document titled The Universal Declaration of Human Rights. The work lists dozens of rights that everyone worldwide is entitled to. Echoing Locke, the United Nations’ list begins stating that “Everyone has the right to life, liberty and the security of person.” It goes on to denounce slavery, torture, arbitrary arrest, and acknowledges rights of marriage, travel, speech, assembly, employment, leisure, food, housing, and health care. What is conspicuously missing from the document is the word “natural,” and in its place we find the word “human.” Rather than claiming that human rights are God-given and grounded in natural law, the document states more modestly that member countries pledge themselves to “a common standard of achievement for all peoples and all nations.” The authority behind this common standard of human rights is the United Nations itself and its member countries. While avoiding the sticky issue of whether rights are “natural,” the United Nations still maintains that human rights are universal and equal—based on the consent of the member countries. Universality and equality is what really matters to us, and not necessarily whether rights are divinely embedded in nature.

**Do We Need Rights?** A second problem for rights theory—natural or human—is whether we really need the concept of “rights” at all. Let’s assume that we have moral duties to refrain from robbing and murdering people. Let’s also assume that we have moral duties to allow people to speak and travel freely. What do we gain by adding that we have rights to these things as well? Political philosophers recognize an interesting relation between rights and duties: the rights of one person entail the moral duties of another. For example, if I have a right not to be robbed, this means that you have a moral duty not to rob from me; if I have a right to speak freely, you then have a moral duty to not interfere with my speaking. This view is called the *correlativity of rights and duties*. The critical question is whether we can just drop all talk of “rights” and make our point using the concept of “moral duty.” Thus, instead of me shouting “Hey, you violated my right not to be robbed!” I might say, “Hey, you violated your duty to not rob me!” Both statements equally suggest that I was morally harmed, and that I am making a claim against the offender. It thus seems redundant to have both the notions of rights and moral duties. If we were to eliminate one in the name of conceptual efficiency, it would be the concept of rights: the language of rights emerged only a few hundred years ago, whereas notions of moral duties have been with us for millennia.
There is, though, an important psychological distinction between duties and rights, which might explain why we place so much emphasis on rights today. Specifically, it’s harder to challenge claims about rights than it is to challenge claims about duties. Suppose that you take my wallet and I shout, “You violated your duty to not rob me!” You’re in a position to dispute my claim since this is a statement about you and the duties that you supposedly hold. You might argue that you’re absolved from your duty because your children are starving and you need to feed them. Even if your argument is flimsy, the burden is nevertheless on me to expose its weakness. On the other hand, if I shout “You violated my right not to be robbed” I can stand fast in my claim since it is a statement about me and the rights that I hold. It makes no difference if your children are starving: I can say that I have a right not to be robbed and that’s that. The difference between rights and duties may only be an issue of rhetorical force insofar as claims about my rights seem more imposing than claims about your duties. But if we’re fighting a tyrannical government, every little bit counts for building an opposition movement, and claims about rights violations may attract more sympathizers. It shouldn’t surprise us, then, that the notion of rights plays such a prominent role during revolutions and when minority groups rise up against social oppression. When the dust from social conflict settles, though, perhaps all that we really have is a system of moral duties that can be expressed with the word “rights” when it suits our purposes.

C. Political Liberalism and Property

Locke influenced a long-standing theory known as political liberalism, which is the view that governments exist mainly to protect individual rights. Foremost among our rights, according to this theory, is the right to private property. Locke himself listed this among his top four rights, and he offers a clear account of how we rightfully acquire property. Suppose that I go to a patch of woods that nobody owns, cut down a tree and carve it into a boat. I am taking some object that is held in common, mixing it with my labor, and thereby making it my property. Once I own the boat, I can keep it, give it away, or sell it as I see fit. If I sell it to someone else, that person then becomes the rightful owner, and I can use the money I’ve made to buy something else. Following this formula, I can accumulate more and more property, all of which I rightfully own. Since property ownership is one of my fundamental rights, part of the government’s task is to protect this right by catching and punishing anyone who tries to steal what is rightfully mine. While Locke’s notion of private property seems reasonable enough, it raises a critical question: can the government tax me—essentially take some of my property—to help people in economic need? The social contract certainly permits governmental taxation for some purposes, particularly to cover the costs of the police and military as they keep the peace and protect my various rights. But can the government rightfully tax me to pay for welfare programs such as those that provide food, housing and financial assistance to the unemployed? The question involves the concept of distributive justice: what is the just way of distributing wealth and poverty in a society? Philosophers within the tradition of political liberalism have vastly different answers to this question.

Nozick and Libertarianism. On one side of the dispute is a theory called libertarianism, which holds that governmental power should be limited to a few basic policing functions. On this view, it’s not part of the government’s job description to pay for welfare programs. A modern champion of libertarianism is American philosopher Robert Nozick (1938-2002) and, like many libertarians, he presents his position within the framework of social contract theory. In the state of nature, Nozick argues, I’m certainly entitled to defend myself against
attacks by others; that task, though, is enormously time consuming, and if I devote all of my efforts to that, there will be little opportunity for me to do much of anything else. Suppose that a salesman approaches me and says, “Good afternoon, sir, I run a small private protection agency in this neighborhood, and for a modest fee I’ll assume the responsibility of defending you against attackers. We’ll not only protect your property, but we’ll track down and punish any violators. Would you be interested in subscribing to this service?” This is exactly the kind of arrangement I need to normalize my life and I’d happily sign up. In fact, the need for this service is so obvious that neighborhood protection agencies would pop up everywhere. To streamline their effectiveness—and prevent war between the agencies themselves—many would band together, creating larger regional agencies. As these consolidate over time, one dominant protection agency would emerge, which would in essence be a bare bones government.

How far should governments expand beyond the role of a dominant protection agency? Not much, Nozick argues. Only a minimal state is justified, and anything more than that will infringe on our rights, particularly our property rights. The manner in which we can accumulate property in this minimal state involves two key principles, which Nozick calls entitlement theory. First, we must initially acquire property by just means—such as by mixing our labor with a commonly held object, as Locke suggests. Second, we must voluntarily transfer that property to another person by just means—such as through a gift or sales contract. No other mechanism of property ownership is available outside of these two principles, such as through theft and fraud. For Nozick, the government itself must respect these principles of property ownership and not reach into our pockets to pay for extraneous projects. For example, if the government taxes me to cover the costs of free housing for the poor, it would be imposing forced labor on me since I’d be working for the advantage of other people without any choice or benefit to myself. While we should sympathize with underprivileged people, any help we give them should be done voluntarily through private charity efforts, and not through governmentally coerced welfare programs.

Nozick offers only one of many possible defenses of libertarianism, but libertarians typically agree that governmental welfare programs are not only unjust for taking our money, but ineffective because they tempt welfare receivers to take advantage of the free ride. Libertarians commonly offer a three-pronged strategy for dealing with poverty. First, libertarians argue, if governments would stop placing restrictions on businesses, the economy would flourish and there would be plenty of jobs for the poor. Second, to protect myself if I ever do become unemployed, I can voluntarily pay into an unemployment insurance program; this would be run by a private company, though, and not by the government. Third, as a last resort we can encourage people to voluntarily donate to charities; these would help only the deserving poor, though, and not lazy free-riders.

As freedom-loving and efficiency-minded as libertarianism is, it faces several obstacles. One is a practical question: would the libertarian solution to poverty actually work? Even without governmental restrictions we can’t always count on a thriving economy. Droughts, population explosions, natural disasters, foreign competition and even bad business decisions routinely throw economies into ruin. The problem can become so great that unemployment insurance and private charities can’t come close to fixing things. Since we look to the government for protection against thieves, murderers and invaders, it makes sense for the government also to protect us from economic disaster, especially when it isn’t our fault.

A second and more critical question about libertarianism is whether its conception of property ownership is fair. [Considering how some people are born with natural advantages over others,] is there
really a fair way of originally acquiring property according to the libertarian scheme? A rival branch of political liberalism called welfare liberalism contends that there isn’t. We turn to that now.

Rawls and Welfare Liberalism. Welfare liberalism is the view that, to address unfair distributions in wealth, the government may tax us to help the needy. Welfare liberalism acknowledges that governments need to protect our rights, including property rights; however, they add, it’s equally important for governments to make sure that some people don’t have an unfair advantage in the race to accumulate wealth. Imagine that an angel was responsible for inserting fresh human souls into the bodies of newborn babies. The angel would reach into a box, randomly pick out a soul, and stick it into the next baby that was born. One soul goes into a baby whose family is living on the streets of Bombay India. The next soul goes into an Eskimo’s baby, and the next into the baby of Bill Gates. Where our souls are placed will almost inevitably determine our financial status in life. The baby in Bombay has no realistic chance of rising above the street-dwelling status of its parents. Bill Gates’s baby, by contrast, will receive private education, go to the best universities, and have an inside track for a very lucrative career. Your financial fate, then, will have been set in motion based on which baby your soul was placed into, and you simply have no say in the matter. While you might like to have your soul placed into Bill Gates’s baby, the odds of that happening are slim, and you can’t claim entitlement to it in any event. American philosopher John Rawls (1921-2002) calls this arbitrary arrangement of wealth the natural lottery, and he argues that justice demands that governments fix the problem rather than perpetuate it.

Like Nozick, Rawls uses social contract theory to make his case. Picture a group of rational, yet self-interested people who gather together to work out the rules of a peaceful society. In this original position, as Rawls calls it, we debate the merits of different systems. Suppose that I’m a homeless person and propose that all of society’s wealth be distributed equally among everyone. Thus, I’d receive $5,000 a year and Bill Gates would also receive $5,000 a year. When we put my scheme to a vote, though, only the poorest people support it while most people who make a decent living oppose it. If most people make over $5,000 a year, there’s no reason why they should support my scheme. We next vote on a plan that a billionaire proposes: everyone keeps all the wealth that they earn, and no money is redistributed among the poor through welfare programs. The richest people in society vote in favor of this while the poorest oppose it. Every time a new scheme is proposed, there is no consensus since each person supports the scheme that benefits him or her the most. It seems that we’ve reached an impasse in our negotiations.

Rawls, though, has a solution: let’s all step behind a veil of ignorance—in essence take an amnesia pill—so we forget what our actual social status is in society. The homeless man doesn’t know that he’s homeless, and the billionaire doesn’t know that he or she’s a billionaire. We then vote once more on our various schemes. Should we divide wealth equally at $5,000 a person? This seems like a bad risk for me since, once the amnesia pill wears off, the odds are slim that my annual income will really be that low. I’ll want some opportunity to make more money if I can, and perhaps even become rich. Should we be allowed to just keep all the wealth that we accumulate? This also seems like a bad risk; while the odds are slim that I’ll be homeless once the amnesia pill wears off, I’ll at least want some financial safety net to catch me if I slip into poverty. The most reasonable scheme we’ll agree to will be a compromise: it will allow financial flexibility for the rich, yet ample protection for the poor.

The precise scheme we’ll accept, according to Rawls, will involve two principles of justice:
First: each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others.

Second: social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone’s advantage, and (b) attached to positions and offices open to all.

Regarding Principle 1, even under the amnesia pill, we’ll all have a basic desire for freedom, and, so, we will want a system that gives us as much liberty as possible, so long as our actions do not violate the liberty of other people. Principle 2 addresses the question of how unequal we should allow people’s wealth to become. Part A of this Principle maintains that the interests of the poor count, and we can redistribute wealth by lowering that of the rich and raising that of the poor. We don’t want to distribute wealth equally, since this will kill personal incentive to work harder. However, in the interest of the poor, we should cap off wealth at the high end. Part B states that everyone should have a realistic shot at the better paying jobs in society, and we should not allow job discrimination, nepotism, or other kinds of favoritism.

The central question for Rawls’s theory is how much taxation should we allow on behalf of the poor? That is, behind the veil of ignorance, how much of a poverty safety net are we willing to create out of the pockets of the rich? Rawls holds that reasonable people will want a lot of economic protection and be willing to place a hefty limit on the wealthy. But Rawls’s critics argue that he is overly-cautious in his efforts to protect the poor. According to his critics, even rational people are gamblers by nature, and, behind the veil of ignorance, I might very well gamble on a scheme that will allow me the greatest financial reward, even if that means putting me at risk of impoverishment. In defense of Rawls, though, while humans certainly are gamblers, we also hedge our bets. We spend large amounts of money on insurance—on our cars, homes and lives—just to feel secure in the off chance that tragedy strikes. This is precisely the reasoning that we’d employ behind the veil of ignorance: risk it all or hedge our bets. It is impossible to know for sure how rationally self-interested people would vote behind the veil of ignorance, but perhaps Rawls is right: we might create a very comfortable financial safety net for ourselves.

D. Individual and Community

In the animated film Antz, a lowly worker ant named Z is in crisis about his personal place in the larger ant colony, which he explains to his therapist:

Z: My mother never had time for me. When you’re the middle child in a family of 5 million, you don’t get any attention. . . . It’s this whole gung ho, superorganism thing that I can’t get. . . . I’m supposed to do everything for the colony, and what about my needs? What about me? I mean, I’ve got to believe that there’s some place out there that’s better than this, otherwise I’ll just curl up in a larval position and weep. The whole system makes me feel insignificant.

Therapist: Excellent! You’ve made a real breakthrough.

Z: I have?

Therapist: Yes, Z. You are insignificant.

Z: I am?

Returning to his job, a fellow ant tries to cheer Z up: “It’s not about you, it’s about us, the team!” Ants are remarkably community-oriented; within their colonies they have complex divisions of labor, all of which focus on sustaining the colony as a whole. Other animals, though, like turtles, have little social structure; except for brief periods of mating, they are on their own from the moment of birth until death. Human beings are paradoxically like both ants and
turtles in our community orientation: while we live in complex communities, we typically have a strong sense of individuality. What should come first, though, the individual or the community?

All of the “liberal” political theories we’ve looked at so far—by Hobbes, Locke, Nozick and Rawls—are individualist: personal liberty is of primary importance, and governments exist to protect us from harms inflicted by others in the community. Even Rawls’s welfare liberalism is individualistic since, behind the veil of ignorance, I’m concerned for the poor only because I myself might be poor. While individualist political theories attract much attention today, rival theories [emphasize how we define ourselves in terms of our communal relations. On these “communitarian”] views, we are first and foremost part of a community [and define ourselves in terms of our communal relations,] so the governments we create should reflect that fact. We will explore two classic community-oriented political theories—those of Plato and Marx.

Plato and the Republic. Imagine a society where men and women live in a totally communal setting. They own no property, receive no pay, eat together in mess halls and sleep in barracks. Men and women have sex with whoever they want, except during a twenty-year period when they are in their prime for reproduction. At that stage, couples are paired for creating offspring with the best traits, and, once born, children are removed from the parents and their identity is kept from them. Children are educated under strict moral supervision and sheltered from degenerate music and literature. There are no squabbles over private possessions or family loyalties; everything is done for the sake of the larger community. While this sounds like the setting of a science fiction movie, it is actually part of Plato’s vision of the perfect society as described in his book The Republic. What makes this especially unusual is that members of Plato’s society recognize that their personal identities are intimately intertwined with others in the community, and it makes no sense to strive for private wealth or power.

What led Plato (427-347 BCE) to propose such a community-oriented conception of society? He explains that societies originally formed to provide for the complex range of needs that we all have. If left to my own devices, I’d have to gather my own food, make my own clothes, build my own house, and learn the countless skills that go into these tasks. We simply can’t do this individually and so, by establishing societies, we specialize in our tasks: I, for example, would make plows, you would grow crops, another would make furniture, and industries of all types would develop to provide us with both necessities and luxuries. In a sophisticated society like this, tensions will invariably develop between our community and other ones on the outside in the competition for natural resources. There’s only so much water, land, and lumber to go around, and we’d like access to these resources just as much as our rivals would. For our society to work effectively and succeed in protecting ourselves from inevitable competition and attack from our rivals, all the parts of our society must work together in perfect harmony. In that sense, we should think of our community like a giant human being, whose physical and mental features must all work together in complete agreement. After all, it wouldn’t work if the mind of this giant person thought “I’ll get a drink of water” while its body grabbed a handful of dirt and ate it.

According to Plato, there are three groups of people in society, which constitute the main parts of this giant person. First there are trades-people—farmers, carpenters, clothiers, merchants—who provide for people’s basic needs. These people are the lowest on the social hierarchy and serve those who are higher up. Second there are guardians who protect society from outside attackers. In fact, the communal lifestyle described above is specifically that of the guardians. They are on the frontlines with warring countries, and nothing should be left to chance in making them the best soldiers possible. Their focus should be on the good of the entire
community, not on themselves. While all classes of society certainly work together for the sake of the larger community, the day to day life of the guardians needs to be infused with community devotion. Third, there are the rulers who decide the best course for society. Selected from among the most outstanding guardians, the rulers constitute the brains and commanding force of the society’s giant person. The top ruler is nothing less than a philosopher-king, who has devoted his life to understanding truth in its purest form and uses this knowledge to direct the entire community.

No society has ever existed along the lines that Plato describes, and Plato himself may not have even intended this to be a blueprint for an actual society. Still, his vision inspired countless depictions of the perfect society, including Thomas More’s famous book *Utopia*. A key question that arises in response to Plato’s community-oriented society is why should we accept our assigned positions? There’s enough individuality in human nature so that we may not feel like participating in a group enterprise like this and instead pursue our private interests. Why am I stuck being a farmer when I prefer the life of a guardian? Plato has an interesting, though somewhat sinister response. We need to trick people into thinking that Mother Nature herself assigned people to their respective social slots. We should convince people that, while all of us are born from the earth, some of us have precious metals mixed into our being. Trades-people are made with a dash of iron and brass, guardians with some silver, and rulers with some gold. The more precious the metal, the higher we are on the social ladder. If this noble lie works, Plato argues, then we should all be content with our social positions, even the humble farmer. In short, sustaining the social hierarchy requires an element of social repression, and this discomforting fact makes Plato’s whole utopian enterprise less than perfect. George Orwell, in his book *1984*, paints a ghastly picture of a community-oriented society which stays together only through governmental lies, intimidation, and brainwashing. The primary challenge of any utopian theory, including Plato’s, is to avoid collapsing into tyranny when reigning in its more independently-minded citizens.

**Karl Marx and Communism.** While Plato may have intended his ideal community to remain just a theory, German philosopher Karl Marx (1818-1883) envisioned a community-oriented society for the real world. He dubbed his theory “communism” and devoted much of his life to spreading its message. In spite of how few people adopted his theory at the time, he felt that global acceptance of communism was inevitable. There are four key ingredients to Marx’s theory. First there is what he calls *historical materialism*. Marx argued that the world is composed entirely of material stuff, with no spiritual component, and that all events mechanically unfold according to rigid laws. Human history in particular emerges very predictably through economic forces. In essence, property and possessions determine everything, and a society’s economic infrastructure is the foundation upon which all other elements of society are built. Second, there is *class struggle*. Throughout history, societies have evolved through conflict between social classes. For example, in ancient times, there was a fundamental class struggle between masters and slaves. This economic system eventually collapsed but it gave way to a new tension in the middle ages between land-owning nobles and serfs who had no choice but to work that land. This conflict was resolved with the creation of a middle class of private business people. In Marx’s own time, he saw a towering class conflict between capitalists who owned industries, and the workers who were virtual slaves to them. He felt that this conflict would also inevitably be resolved—with the creation of communism—but getting there was the trick.

The third element of Marx’s theory is *alienated labor*. We might ask Marx, what’s so bad about rich industrialists providing jobs for
factory workers who need the money? It seems like a reasonable economic relationship. Marx’s answer is that human beings forge their identities through labor and, by working for the rich industrialist, a worker’s identity is fractured. As hard as we labor, it’s all for the company, in exchange for which we receive a measly paycheck. It’s not just that we’re underpaid for our hard work, or that our jobs are torturously monotonous as we perform the same task hour after hour, and day after day. As bad as these things are, the real insult is that we’re forced to hand a piece of our identities over to our bosses, just to survive. It’s much like what a prostitute does, Marx argues, insofar as she sells off the most intimate part of her being in exchange for some financial security. We are alienated from our labor, and, for that reason, alienated from a part of ourselves. Everything within a capitalist society supports this alienation—the laws, the government, and even art. We see this most clearly in religion, though, where we’re taught that God requires that we suffer on earth through our miserable jobs, for which we’ll be rewarded in the afterlife.

The fourth component of Marx’s theory is communist revolution. He argues that when enough workers are fed up with their oppressive conditions, they will launch a communist revolution, the aim of which is to abolish private property. In some countries the revolution will be catastrophic and bloody, since rich property owners won’t willingly hand their possessions over to the masses. In others, it will be progressive, where a growing number of workers eventually overtake the crumbling capitalist infrastructure. In The Communist Manifesto (1748) Marx explains that, along with the elimination of private property, the state will act on behalf of workers and take control of the country’s economic resources. Education and other essential social services will be free. In time, class distinctions within society will disappear; governments, which historically have existed to oppress the working class, will no longer be needed.

Marx’s theory might have remained a 19th century curiosity were it not for the Russian revolution in 1917 which embraced communism wholeheartedly, and transformed much of 20th century political ideology. Critics of communism throughout the 20th century have questioned whether humans were capable of following a community-oriented economic system. Personal greed is an important element of social progress, they argue; we work harder and are more creative when the prospects of wealth and power are dangled before us. While it’s nice to think that I’d work my hardest on behalf of the larger community, I don’t seem to be designed that way. As proof of this point, critics today often draw attention to the widespread collapse of communist governments in the late 1980s: if Marx was right, why have these countries abandon communism in favor of capitalism? Marx recognized that people have an enormous capacity to be selfish; still, he argued, we are indeed designed to be community-oriented. We have a special human quality, which he dubbed our species-being, which reflects our more evolved human nature and prompts us to see ourselves as part of a collective whole. This, Marx thinks, is what will drive all people in the future. The issue between Marx and his critics, then, is whether something like species-being is embedded in our notions of personal identity.

E. Governmental Coercion

Governments have a difficult task of deciding which kinds of actions should be legal or illegal. Lawmakers in Virginia considered imposing a $50 fine on people wearing low riding pants that expose their underwear. They abandoned the idea, though, when it drew public ridicule. A man from New York was arrested for telling lawyer jokes in a courthouse. For years he had used confrontational devices like this to push for greater public access to the courts. In this situation he was accused of being abusive and causing a disturbance; the charges, though, were dropped for lack of evidence. A man from New Mexico was charged with distributing sexually oriented material to minors because of cartoon
bumper stickers on his car that depicted bare-breasted female devils in sexually provocative positions. The man said “I’m offended by church people saying I can’t drink on Sundays, so I put the devil chicks on my car, because I figured it would offend them right back.” He stated that police were charging him with this crime only because “some overzealous, churchgoing detective got offended by it, and got even more offended by it by the fact that I didn’t take it off after he threatened me.”

All three of these examples are startling since they illustrate the government’s tendency to restrict our freedom more than we’d prefer. Governments are by their very nature coercive. In their efforts to keep the peace, they must set boundaries for our conduct and punish us when we cross those lines. We understand why governments prohibit theft and murder, and we’d be outraged if our laws permitted serious offenses like these. At the same time, though, we expect the government to allow us freedom to act as we individually see fit—such as wearing low riding pants or making lawyer jokes in a court house. One task of political philosophy is to examine general guidelines for determining when governments may restrict our conduct.

Four Principles. Governmental restrictions are often justified on four distinct grounds. The first and most important is the harm principle, which is that governments may restrict our conduct when it harms other people. Laws against assault and murder are clear examples of this. But the notion of “harm” is a little fuzzy, and we’re faced with an important question: what counts as harm? When we think about harm, physical attack is what immediately comes to mind. What is central to the notion of harm, though, is that it is an invasion of an interest, and avoiding physical injury is just one of these. I also have an interest in my property, and so I’m harmed when someone steals from me. My interest in my personal reputation means that I’m harmed when someone slanders me. My interest in privacy means that I’m harmed when you trespass on my property or peek through my window. We have so many interests, though, that what counts as “harm” can quickly get out of hand. I have an interest in emotional happiness, so should there be laws preventing people from telling me unhappy news, such as news about a car wreck down the road? But key to the notion of harm is that it involves a serious injury to an interest, and not merely a trivial injury. Hearing bad news is part of life, and even if it makes me sad to hear about a car accident, there is still something very commonplace about this, and it would not count as a “harm” to my emotional happiness. So, while our common concept of harm is vague, it helps to think of it as a serious injury to an interest.

Second is the offense principle, which is that governments may keep us from offending others. Virginia lawmakers, for example, attempted to prohibit low-riding pants on the grounds that the public display of one’s underwear is offensive for many people. Flag burning, pornography, foul language and even lawyer jokes might be restricted on the grounds of offense. Like the notion of “harm,” what qualifies as an “offence” may quickly escalate. Suppose that I’m on a park bench and a man sits down next to me. He smells horribly, starts picking his nose, and then sings “Who Let the Dogs Out” at the top of his lungs. I try to ignore him but eventually I’m forced to walk away. As bothersome as this might be for me, it certainly wouldn’t require legal intervention, and the whole experience would be more like a nuisance rather than a full blown offense. Offenses, properly speaking, involve a more intense reaction of outrage, and they typically involve situations that I can’t avoid without major inconvenience. My park bench experience didn’t outrage me, and it was easy enough to remove myself from the situation.

Third is the principle of legal paternalism, which is that governments should prevent people from harming themselves. While the harm principle, noted above, focuses on the harm that I might cause other people, the principle of legal paternalism
involves the harm that I might do to myself. As such, the government would restrict my choices for my own good. The term “paternalism”—literally “father-ism”—means treating people in the way that a father treats a child. The very notion is degrading since it presumes that we behave like reckless children and need the government to save us from ourselves. Sometimes, though, this is indeed an accurate description of our behavior. Sports such as bare knuckle boxing are outlawed for this reason, and others like football are heavily regulated to minimize harm to participants. Laws requiring seat belts, motorcycle helmets, and doctors’ prescriptions to purchase drugs all aim at preventing harm through our carelessness, stupidity, or desire to take dangerous risks. There is quite a lot of harm that I can voluntarily do to myself and, according to paternalists, society should prevent that harm if it has a chance.

Fourth is the principle of legal moralism, which is that governments may restrict conduct that is especially sinful or immoral. While harmful acts such as stealing are certainly immoral, this principle focuses on actions that don’t necessarily cause harm. What is at issue is the moral or religious well-being of a society, and the need to stop people who undermine this. For example, in religiously conservative Saudi Arabia, celebration of Valentine’s Day is banned and, along with that, the sale of red flowers. The group responsible for enforcing this and similar morally-based laws is “The Committee for the Promotion of Virtue and Prevention of Vice.” Even in the much more liberal country of Great Britain, their laws prohibit publishing religiously blasphemous material. Legal moralism is the justification behind many sex laws, such as those prohibiting homosexuality and oral sex, even when these acts are done privately between consenting adults. Legal moralism is most prevalent in countries that have a moral or religious code that most citizens share. Muslim countries like Saudi Arabia, which are governed by Islamic law, are good examples. But in more pluralistic countries such as the United States, it is more difficult to find a core set of values upon which everyone agrees, so laws grounded in legal moralism, such as those against homosexuality, often look like one group ganging up on another.

**Mill’s Principle of Liberty and Harm.** The central point of dispute surrounding the above four principles involves determining which of them are valid reasons for governmental coercion. In perhaps the most famous discussion of this subject, British philosopher **John Stuart Mill** (1806-1873) argues that only the first principle—the harm principle—is justified. He makes his point here:

> [T]he sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. **[On Liberty, Chapter 1]**

Mill states here that individual liberty should only be restricted when our actions harm others, but not when they simply harm ourselves. Thus, he rejects the principle of legal paternalism. Elsewhere in this work he criticizes religious groups that seek to restrict behavior on moral grounds or because it offends. Thus, he also rejects the offence principle and principle of legal moralism. For Mill, this means that I am free to harm myself, behave immorally and offend others, so long as I don’t harm anyone else. It would be difficult to articulate a concept of personal liberty any stronger than this, and many people think Mill has gone a little too far.

Why would he propose such an extreme view? Part of his motivation undoubtedly owed to his own experience as an open-
minded person in a traditional 19th-century society. He held liberal views about God, the equality of women, the social role of government, and other issues which conservative critics would have liked to suppress. In *On Liberty* he explains why it is so important to allow people a zone of free expression. Each of us, Mill argues, understands his or her own interests better than any well-intentioned outsider. We spend much of our time creating a lifestyle that works for us individually. There is no one-size-fits-all way of life, and it is typically through trial and error that I discover what’s best for me. This involves making mistakes—and sometimes mistakes that harm me. Nevertheless, Mill argues, society will be a better place if we’re allowed to make decisions on our own without the intrusion of others. In a nutshell, his point is that a wide sphere of personal liberty is essential for a happy society.

Mill’s “happy society” argument is just one way of defending the harm principle, but a more common approach is based on social contract theory. In the state of nature, we have complete liberty to behave as we please; the price we pay for this absolute freedom, though, is brutal war with everyone. To end the state of war, we’re willing to give up some of our freedom, but no more than is absolutely required for achieving peace. To attain peace, all that we really need to do is mutually agree to avoid harming each other; that is, we should adopt the harm principle, which will permit the government to punish me if I harm others. Peace will not come about any more quickly by also adopting the principles of offence, legal paternalism, and legal moralism. Rather than having the government punish us for offensive, self-harming, or immoral conduct, we can simply be tolerant and live peaceably with each other. Embracing these three extra principles, then, is an unjustified restriction on our freedom.

The idea of a maximally free society is certainly appealing; we’d all like the opportunity to at least occasionally do odd-ball things without the government clamping down on us. The “happy society” and social contract arguments each attempt to defend this intuition in a different way. As compelling as these two arguments are, though, societies are reluctant about completely throwing aside the principles of offence, legal paternalism, and legal moralism. Take the principle of legal paternalism, for example. Assume Mill is correct that society is better off when we have the freedom to explore possibilities. Still, society might be slightly more happy if governments paternalistically blocked us from our most stupid and harmful actions. This might be particularly so for people who continually make bad choices for themselves by gambling their income away, being in a continual drunken stupor, eating rotten food from dumpsters, or living in a shack that’s about to collapse. Suppose also that social contract theories are right that we can achieve a peaceful society through the harm principle alone. We need to remember, though, that the reason I sign the social contract to begin with is to make my life safer; in Hobbes’s words, I want something better than a life that is “solitary, poor, nasty, brutish and short.” The principle of legal paternalism aims exactly at making our lives safer—essentially by keeping us from attacking ourselves. So it makes sense for signers of the social contract to agree to some principle of paternalism along with the principle of harm.

F. War

While the Interim Government of the Republic of Texas has not yet declared war on the United States, they came close to doing so several years back when their leader threatened war if authorities attempted to arrest him. “Once they make the move,” he stated, “we’ll have millions of Americans on our side—including every militia in the country. We’re talking war here.” Had authorities arrested him, the truth of the matter is that very few Americans would have come to his defense. Imagine instead, though, that their leader was right and events unfolded like this. Once arrested, millions of Americans sympathize with him, just as he predicted.
Paramilitary and survivalist groups around the country load up their pickup trucks with weapons they’ve been secretly stockpiling for years, and head down to Texas to take a stand. Declaring war in a campaign they call “Texan Freedom” they ambush police, free their leader from jail, and roll into Austin taking control of the capital. Initially caught by surprise, the U.S. military then descends on Texas, attacking rebel strongholds and arresting sympathizers. In spite of mass casualties, the rebels retain control. The war abruptly ends when a nuclear bomb explodes in Austin—each side accusing the other of detonating it.

Would this war of “Texan Freedom” be a morally justified one? That is, are there specific features of this war that show decisively whether it was either on the side of right or the side of wrong? Since the middle ages, philosophers have weighed in on the subject of war, particularly on the moral question of when, if ever, wars are permissible. Two traditions are central to the debate. First is just war theory, which maintains that wars are sometimes morally justifiable. Second is pacifism, which is the view that wars are never morally justifiable.

Just War Theory. In Western Civilization, just war theory has been the dominant philosophical view of war since the Roman Empire was threatened by barbarian invaders. While the earliest Christians were largely pacifists, Saint Augustine argued that Christians could with clear conscience answer the Empire’s call to military duty. Following Augustine’s lead, philosophers over the centuries have developed specific criteria for determining when a war is or is not justified. Just war theorists distinguish between two issues: first, the moral requirements for initially waging a war, and, second, the moral requirements for conducting a war once it’s been started.

The issue of initially waging war is traditionally designated by the Latin phrase *jus ad bellum*, which literally means “law to war.” While just war theorists have offered different criteria for waging war, four points are central. First and most important is that there must be *just cause* for waging war, and the most crucial of all causes is resistance of aggression. The aggression must be quite serious in nature, such as physical violence through bombing campaigns or military invasion, and threaten an entire population’s rights. When Japan bombed Pearl Harbor, their aggression was serious enough to justify the U.S. in declaring war. Does the war of “Texan Freedom” described above have just cause? It may very well be that the Interim Government of the Republic of Texas has a legitimate complaint against the U.S. government for continually limiting the liberties of its citizens by increasing taxes or limiting gun ownership. Libertarians have been arguing as much for years. Nevertheless, even if this complaint is legitimate, the U.S. government cannot be accused of anything like serious physical aggression that threatens all Texans. And, without the threat of serious aggression, the war of “Texan Freedom” would be without just cause.

The second criterion is right intention. The motive behind waging war must be proper, such as wanting to return to the state of peace prior to the aggression. Wrong intentions would be the desire to acquire land, to plunder the resources of another country, to take vengeance, or to vent racial hatred. One motive behind the war of “Texan Freedom” is a desire for liberty, which on face value seems legitimate. However, behind their campaign is a longing for the good old days when Texas had a more independent identity as a sovereign country. Nostalgia, though, does not qualify as right intention. It is normal for people to have some feelings of local pride, but the danger is when this expands into a zealous nationalism which gives countries a false sense of superiority and motivates them to impose their will on their neighbors. Japan was seduced by this zeal in World War II, which prompted them to seize control of surrounding countries. Determining right intention sometimes involves hunting down hidden motives, such as
nationalism, which lurk beneath more noble ones like the desire for liberty.

The third criterion is **proper authority**: the war must be publicly announced by the legitimate authority and made known to the enemy. The U.S. Government was the proper authority for declaring war on Japan after the bombing of Pearl Harbor, and it did so with a public declaration. With the war of “Texan Freedom,” we need to consider the authority behind the Interim Government of the Republic of Texas. Who do they really speak for? If they had the support of most Texans, then perhaps they would be legitimate. On the other hand, if their support came principally from their own members and sympathetic militia groups, then they would not have proper authority.

The fourth criterion is **reasonable success**: the war shouldn’t be pointless, since it is wrong to sacrifice human lives and squander economic resources if the outcome is unlikely. Determining the odds of victory is often difficult since it involves complex calculations about weaponry, support from allied countries, fuel supplies, and unpredictable factors such as the weather. At least sometimes, though, the likely outcome is clear. In World War II, France realized that declaring war against Nazi Germany would be futile, so they reluctantly accepted German occupation. Any effort like the war of “Texan Freedom” would similarly be doomed to failure and would thus not have a prospect of reasonable success. All in all, the war of “Texan Freedom” does not fulfill any of the four criteria of initially waging a just war.

Let’s turn now to the other issue of just war theory, namely the moral rules of **conducting** a war; this is designated by the Latin *jus in bello*, literally “law in war.” The point is that neither side of the conflict can use reprehensible techniques in achieving its cause, regardless of how just that cause is. There are two criteria here. First is **discrimination**: both sides of the conflict must identify legitimate targets. War should aim at the people responsible for the wrong, not the innocent; this means that soldiers can’t target civilians in residential neighborhoods. Some civilian deaths are inevitable in war, but they should be minimized, and, in any event, not be intentional targets. Wars of independence, like “Texan Freedom,” are especially troubling for blurring the distinction between combatants and civilians. What starts out as combat between soldiers may quickly escalate to the genocide of sympathetic civilians. The second criterion is **proportionality**: the military should only use the amount of force that is required to achieve their goal. The aim here is to minimize war’s destruction. Weapons of mass destruction typically go beyond such goals and thus their use would be unjust. The more civilian deaths, the more questionable the justness of the war appears. In the war of “Texan Freedom,” the use of a nuclear weapon clearly went too far regardless of which side was responsible for launching it.

The principles of just war outlined above are theoretical—devised by philosophers for assessing the more abstract problem of the morality of war. In addition to these theoretical discussions, though, there have been historical agreements worked out by governments that establish codes of conduct for actual military forces. The most famous of these is the Geneva Convention, which is a series of international treaties that regulate the treatment of civilians in occupied countries and prisoners of war.

**Pacifism.** While just war theory has predominated among philosophers since the days of Augustine, there have continually been pacifist groups opposing war. Pacifists certainly recognize the dangers of aggressive countries, such as Germany and Japan in World War II. They also recognize the importance of taking action against aggressors, such as through non-violent resistance, economic boycotts and shunning by the international community. What they oppose, though, is a campaign of killing. Some pacifists—called **absolute pacifists**—feel that all wars, with no
exception, are wrong. Others—called conditional pacifists—object to wars in principle but feel that some are permissible in extreme emergencies. It is the position of absolute pacifism that is more philosophically interesting, which we will explore here.

Considering how crucial some wars have been in stopping aggression, such as World War II, why would someone hold to absolute pacifism? The most common justification is religious pacifism: war is contrary to religious teachings. Most religious scriptures from around the world contain some general condemnation of killing, such as “Thou shalt not kill” from the Ten Commandments. Some theologians have interpreted these injunctions literally as condemning any intentional killing, even in self-defense or in times of war. Further, many religious founders, such as Buddha and Jesus, conducted their lives non-violently, and some believers feel that they represent a model of pacifism which we all should adopt. Religious pacifism, though, is largely a theological question that hinges on subtle interpretations of a faith tradition. It will thus be compelling only for believers who adopt a specifically pacifistic understanding of their faith.

There are, though, more secular grounds for pacifism. One argument is that the benefits of war never outweigh its costs. This is particularly evident in modern warfare which is especially destructive and doesn’t seem to serve the long-term interests of society. The threat of nuclear war is the best example of this. If the only way to subdue an enemy is to destroy much of the world with nuclear weapons, it’s hard to see how society would benefit from engaging in war in the first place. While this cost-benefit argument may compel us to morally condemn nuclear war, it might be different with more modest wars. This would be so if by declaring war we can quickly subdue an aggressor using only conventional weapons and causing only a minimal number of deaths. The cost-benefit argument, then, may only apply to the more extreme types of war, and not to all wars.

A second and more compelling secular argument is that war violates our foundational duty to avoid killing innocent people. This is a basic moral obligation that all civilized countries recognize, irrespective of religious traditions. Now, in modern warfare we can count on enormous numbers of innocent civilian deaths—perhaps even more than the numbers of soldiers that die. No matter how hard a country tries to restrict casualties to soldiers, it invariably spills over into the civilian population. While it may be difficult to precisely determine what a morally acceptable level of civilian deaths might be in an otherwise well-intentioned war, modern wars have almost certainly gone beyond that level.

While these arguments for pacifism may not be 100% convincing, they nevertheless show that the pacifist position has some reasonable foundation and that pacifists cannot be simply dismissed as irrational. But just war theorists respond to pacifism with a common criticism: pacifists are free-riders. In spite of their noble talk about the evils of killing, pacifist themselves enjoy the benefits of a protected society without participating in its defense. Thus they are reaping the benefits of war and military defense without paying the costs. This criticism is unfair, though, since pacifists are not given the opportunity to try out their own non-violent solutions. They are forced to live in a society controlled by warmongers who continually reject more peaceful solutions. It’s possible that the pacifist’s solution would be substantially better for society than the warmonger’s. In that case, pacifists—and everyone else—end up suffering from the decision to go to war, and by no means benefit from it. Pacifists are thus not free-riders: it’s more like they’ve been overpowered by leaders who are devoted to war.

War is an unfortunate concern for any government, and the history of human civilization is connected by a string of bloody conflicts, one after the other. Political philosophers from Plato to Hobbes to Nozick have argued that societies were in fact first formed as a means of protection from inevitable outside attack. Maybe wars
will end in the distant future, but for the time being they are a fact of life for virtually all countries around the world. While the time for pacifism may not yet be here, it still makes good sense to think hard about the justness of military conflict before declaring war on an enemy.

Throughout this chapter the running theme of political philosophy as been protection: it is the principal justification of a government’s existence, and it remains its fundamental task. We expect governments to protect us against criminals, rights violations, economic injustice, exploitation, poverty, offense, our own personal stupidity, and foreign invasion. At the same time, though, we have a strong pull towards individuality and personal freedom, which often conflicts with government efforts to protect us. Is there a happy middle ground? One of the more unique features about political philosophy is that it doesn’t search for absolute truth as much as other branches of philosophy do. With philosophy of religion, for example, we want to know as a matter of fact whether God exists. In the realm of politics, though, there is great latitude regarding the kinds of governments that human beings can establish. This changes over time, and the larger question becomes what kind of political systems are most suitable in the present state of affairs. Thus, in attempting to find the right balance between government protection and personal freedom, a one-size-fits-all solution may not be realistic. The best we can do is consider whether, at this particular point in history, our society could benefit from either more protection or more freedom with the myriad of issues that we face. Should our government provide universal health care, or make college education more affordable, or force business to protect the environment, or restrict gun ownership? When we do attempt to find a balance between protection and freedom on these issues, the theories we’ve examined in this chapter will inevitably play an important role.

For Review

1. Describe Hobbes’s state of nature and list his laws of nature.
2. How does Hobbes’s state of nature parallel the prisoner’s dilemma?
3. Describe Locke’s account of natural rights and how it may justify revolution.
4. Explain Bentham’s criticism that natural rights are not grounded in fact.
5. What is Nozick’s theory of the minimal state, and what does this imply about taxation to help the needy?
6. What is Rawls’s theory of the veil of ignorance, and what does this imply about taxation to help the needy?
7. Describe the three classes of Plato’s perfect society, and what governments must do to keep people content with their assigned social class.
8. What are the four key elements of Marx’s theory?
9. Explain the four grounds of governmental coercion.
10. What is Mill’s view of liberty and harm, and what is his “happy society” argument?
11. Explain the four criteria of waging a just war, and two criteria for conducting a just war.
12. What are the religions and secular arguments for pacifism?

For Analysis

1. One criticism of social contract theory is that it allows for the possibility of a dominant group subjugating minorities. How might a social contract theorist respond to this criticism?
2. Write a dialogue between Locke and Bentham on the subject of natural rights.
3. Write a dialogue between Nozick and Rawls on the subject of welfare taxation.
4. Criticize Marx’s view of communism from Locke’s perspective of natural rights.
5. Defend the principle of legal paternalism against Mill’s “happy society” argument.
6. Write a dialogue between a just war theorist and a pacifist on the subject of the moral permissibility of war.

REFERENCES AND FURTHER READING

Works Cited in Order of Appearance


Further Reading