This is a book about frames and framing effects. So, let’s start with some examples, to get a preliminary feel for what is at stake here. I will present five framing effects, without much by way of discussion or analysis. In each of these cases, people come to view a single outcome in very different ways, depending on how it is framed.

The first three cases are experimentally induced in a lab. They display the framing effects cleverly and very clearly, even though they are not the most celebrated cases of experimental framing effects such as the celebrated Asian disease paradigm, which we’ll look at in more detail in Chapter 2.

The fourth example shifts the emphasis away from the laboratory. It is an idealized version of familiar real-life cases with which most of us are painfully familiar. When faced with temptation, decision-makers struggle to stick to the plans they have committed themselves to as the moment of truth draws near. We’ll look at cases of temptation and self-control in much more detail later on (in Chapter 7), but the point I want to make here is that how we frame the path of virtue (or the path of temptation) can determine whether or not we manage to exercise self-control.

The final example completely changes tack. We go back to the shadowy world of the ancient Greek tragedies, where history and myth blend. It is a famous passage from the first play in Aeschylus’s trilogy *The Oresteia*, where the chorus looks back to Agamemnon’s fateful dilemma at Aulis. From a psychological point of view, it is exponentially more complex than either the three experimental cases or the self-control/temptation example (and certainly doesn’t lend itself to experimental replication). Despite that (or really, because of it), we will come back to it many times in the course
of this book, as I believe that the power and importance of frames becomes much clearer when we see how they function in the really hard cases. The easy cases have received too much attention, which has skewed our understanding of frames and framing effects.

Without digressing too much, at this stage the only goal is to understand the cases and come to a (perhaps provisional!) conclusion about what they reveal, and whether what is going on is rational or irrational. It might be a good idea to make a note of your thoughts, so that you can come back to your immediate reactions to these examples when we are further along in the book.

**Framing Effect 1: Rating Basketball Players**

Irwin Levin at the University of Iowa asked subjects to evaluate how well basketball players were performing, based on information he provided about their shots over a period of time.¹ He presented the information in two different ways. For one group, it was presented positively, as they were told the percentage of shots that the player made successfully. For a second group the information was presented negatively, in terms of the percentage of shots that the player had missed. These are of course different ways of framing the same facts about how the player played. Yet Levin found that the same players were consistently ranked more highly by subjects in the positive frame than they were in the negative frame.

**Framing Effect 2: Negotiating Contracts**

Margaret Neale and Max Bazerman (from Arizona and MIT respectively) asked a class of 102 undergraduates studying Business Administration at the University of Texas to simulate an industrial negotiating situation.² The students had to imagine that they were negotiating with union representatives on behalf of a fictional company (Townsford). Their job was to negotiate a settlement, but they also had the option of giving up and going into binding arbitration – a much riskier strategy.

They were divided into two groups. Both groups were presented with information about the different priorities and settlement-points of management and the union on five different issues. For one group the information was presented positively (from the perspective of the company).

Students in the positive frame were given numbers corresponding to the total gain to company if the company were to settle at that point. They were also told:

Any union concession from their current position will result in gains for the company. Please remember that your primary objective is to maximize such gains for the company. I cannot emphasize the importance of these gains to Townsford enough. It is mandatory that you, as Townsford’s representative, secure such concessions from the union to increase these gains to a meaningful level.

Students in the negative frame were given exactly the same numbers, but those numbers were presented in the form of losses rather than gains. This group was told:

Any concessions beyond those granted will represent serious financial losses to the company. Please remember that your primary objective is to minimize such losses to the company. I cannot emphasize the severity of this situation enough. It is mandatory that you, as Townsford’s representative, secure the necessary concessions from the union to reduce our losses to a tolerable level.

Obviously, there is no difference in the objective information possessed by the two groups. Contract negotiations are what is called a zero-sum game. A gain to the company is a loss for the union, and a gain to the union is a loss for the company.

Still the group in the positive frame were much more likely to negotiate a settlement, whereas the group in the negative frame were more willing to take the riskier option of abandoning the negotiation and submitting to a binding arbitration.

Framing Effect 3: Sacrifices for the Common Good

As game theorists know well in theory, and the rest of us in practice, many social situations have the form of a social dilemma. Social dilemmas occur when collective disaster is the result of individuals behaving perfectly rationally to promote their self-interest. Open range grazing in the American West is a famous example. It had its heyday in the second half of the nineteenth century. Any rancher could graze their animals on open rangeland and each individual farmer had an obvious incentive to put as many of their animals on the land as they could. Why not, since the grazing is free? But of course, if too many farmers do so, then the

---

rangeland is destroyed for everyone. Open range grazing is an example of a commons dilemma (often called tragedy of the commons). In the simplest form of commons dilemma individuals have to decide what share to take for themselves of a shared resource (as in the open range case, where the open range is the shared resource).

Another type of social dilemma comes with the provision of public goods. A public good is a good that benefits everyone, at least potentially (such as university education, or state-funded healthcare in a single-payer system, such as the United Kingdom’s National Health Service). In public good dilemmas, individuals have to decide how much (if anything) to contribute to maintaining a public good. Such dilemmas can arise for private groups – residents of an apartment block deciding whether to increase the maintenance charge to pay for a new roof, for example. But they also arise in debates about taxation levels. In the United States, for example, local governments sometimes hold referendums on increasing property taxes to pay for additional local services, or improved schooling. Each resident and each voter is confronted with a public good dilemma. Marilyn Brewer (UCLA) and Roderick Kramer (Stanford) ran a study to test whether subjects would respond differently to a collective choice problem depending on whether it was presented as a commons dilemma or as a public good dilemma. The experimental task was cleverly designed to induce the tension between individual good and common good that characterizes all social dilemmas. Subjects were told that there was a common resource pool of points. All the subjects had access to the common pool and were instructed to maximize their own points total while maintaining the common resource as long as possible. In one condition (the public good condition), subjects were given points and then had to decide what proportion to contribute back to the pool, while in the commons dilemma condition subjects had to decide how many points to take from the common pool. The outcomes were identical across the two conditions in terms of points. And so the monetary rewards to the subjects

\footnote{Limited forms of open range grazing persist in some western states in the United States and Canada, but when a state such as Texas is described as an open range state, what this typically means is that landowners do not have a legal obligation to fence their animals and, for example, keep them off public roads. The “golden age” of open range grazing was brought to an end in the United States in the last years of the nineteenth century by a combination of over-supply and over-grazing, compounded by a very severe winter in 1886–87.}

\footnote{The phrase originated with the Victorian economist William Forster Lloyd, but was popularized by the ecologist Garrett Hardin in an influential paper of the same name published in 1968.}
were the same. There is no difference, for example, between starting with 1,000 points and contributing 250 points to the common pool, on the one hand, and starting with 500 points and taking 250 points from the common pool. Either way you end up with 750 points.

Still, the two groups behaved very differently. It turns out the subjects left more points in the common pool in the commons dilemma condition than they were prepared to contribute to the common pool in the public good condition. Apparently, people are much less willing to contribute points to the common pool than they are to leave points in the common pool. Forgoing a gain is easier than taking a loss, it seems, even when the experiment is designed so that there is no difference in outcome, but only a difference in how the outcome is framed.

**Framing Effect 4: The Battle against Temptation**

It is easy to make commitments in advance, but hard to live up to them when the time comes to follow through in the face of temptation. The basic phenomenon should be familiar to anyone who has taken out a gym membership or made a New Year’s resolution to lose weight. At a safe (temporal) distance the long-term outcome of being fit and slim is far more attractive than the short-term prospect of an extra hour in bed, or the mid-morning snack. And yet when the alarm goes off or the stomach starts rumbling hours after breakfast and hours before lunch it is a different story. The immediate reward suddenly seems far more attractive than the long-term outcome. Self-control is hard. In fact, one might wonder how it is even possible.

We need to exercise self-control because preferences change over time. In the indeterminate future, being fit and slim is much more attractive than the prospect of a snack. But when the snack is right there, it seems much more appealing than being fit and slim at some indefinite time in the future. This type of preference reversal occurs because of how people discount the future. If I have a high discount rate, then I care relatively little about the future. But if I have a low discount rate, then I care very much about what happens in the future. The problem is that people do not typically have constant discount rates. Much experimental evidence suggests that the discount rate for a given event changes as the event approaches. If I decide on a Friday to fast until lunchtime on the following Thursday, then I probably have a high discount rate on Friday and over the weekend for the breakfast that I am planning to forgo on Thursday morning. As the week goes on, though, my discount rate for the breakfast...
falls. And by the time I wake up on Thursday morning it is very low indeed. In the meantime, though, my discount rate for the long-term goal of being slim and fit has not really changed at all. And so the short-term prospect of eating breakfast comes to seem more important than the long-term goal of being slim and fit. That is how temptation strikes!

Sometimes we succumb to temptation. But often we don’t. Why not? There is a vast literature on this, from self-help manuals to experimental studies on how rats respond to delayed rewards. Crucially, though, whether we succeed in exercising self-control can be due to how we frame the different possible actions and outcomes when faced with temptation. If it is a simple choice between eating breakfast and sticking to my fasting plan, and if my changing discount rates have led to a preference reversal, then I may well end up chowing down on my breakfast. But what if I attach a special importance to actively resisting temptation? This might lead me to a different way of framing the act of holding out for the long-term reward of being fit and slim. For example, if I frame it as the act of being resolute – and I like the idea of being resolute – then it fits with my self-conception. And being resolute now may well make it more likely that I’ll be resolute in the future. For all these reasons I might well prefer being resolute in the face of temptation to having breakfast, especially if I frame having breakfast as succumbing to temptation.

As we’ll see in Chapter 7, there is experimental evidence that self-control often works like this. But really this is a framing effect. All I’ve done is reframe the outcomes and reconceptualize the decision problem. In this situation there is no difference between being resolute in the face of temptation and forgoing breakfast. And succumbing to temptation is the same as eating breakfast. The outcomes are the same. Only the framing changes.

Hopefully, by this point you will have started to wonder whether framing effects are always irrational. On the face of it, self-control is a good thing. In fact, it seems more irrational to succumb to temptation and abandon a long-term plan. So, it seems odd to make it irrational to escape temptation by reframing outcomes. This is a case where framing seems to be a tool for rational thought and rational action.

**Framing Effect 5: Agamemnon at Aulis**

And now for something completely different. The last framing effect I want to present comes from Greek tragedy. As I mentioned earlier, it sits far away from the experimental studies of framing discussed up to
now and, although I will come back to it in much more detail in Chapter 6, I want to put it on the table now to introduce some of the complexities and richness of the framing phenomenon.

The chorus in Aeschylus’s tragedy *Agamemnon*, the first play in the *Oresteia* trilogy, tells the story (familiar to his audience from many other sources) of the Greek leader Agamemnon at Aulis. Agamemnon is leading the Greek fleet against Troy to avenge the abduction of Helen by Paris. While the fleet is becalmed at Aulis, the prophet Calchas interprets a portent – two eagles swooping down to kill and eat a pregnant hare. As Calchas interprets the portent, it reflects the displeasure of the goddess Artemis at the prospect of innocents being killed at Troy. The lack of wind has the same source. The only solution, says Calchas, is for Agamemnon to sacrifice to the goddess his own daughter Iphigenia.

In a powerful and memorable passage, the chorus recalls Agamemnon’s anguished cry:

> And I can still hear the older warlord saying,
> “Obey, obey, or a heavy doom will crush me! –
> Oh but doom will crush me
> once I rend my child,
> the glory of my house –
> a father’s hands are stained,
> blood of a young girl streaks the altar.
> Pain both ways and what is worse?
> Desert the fleets, fail the alliance?
> No, but stop the winds with a virgin’s
> blood,
> feed their lust, their fury? – feed their fury! –
> Law is law! –
> Let all go well.”

With apologies to Aeschylus (excellently translated by Robert Fagles), Agamemnon might more prosaically be described as in the grip of a framing effect. There is a single option, bringing about the death of Iphigenia, that Agamemnon frames in two different ways – as *Murdering his Daughter*, on the one hand, and as *Following Artemis’s Will*, on the other. His alternative is *Failing his Ships and People* (by refusing to make the sacrifice).

Agamemnon’s dilemma is that he evaluates the death of Iphigenia differently, depending on how it is framed. He certainly prefers

---

Following Artemis’s Will to Failing his Ships and People. At the same time, though, he prefers Failing his Ships and People to Murdering his Daughter. But he knows, of course, that Following Artemis’s Will and Murdering his Daughter are the same outcome, differently framed.

By way of a taster for what lies ahead, my view is that the last two examples of framing effects (the battle against temptation and Agamemnon at Aulis) are fundamentally different from the first three. They are more complex both because the decision-situations are more multifaceted and because they engage reasoners’ motivations, emotions, and values in deeper ways. It is here that we need to look properly to understand the power of frames; to see how there can be rational framing effects; and to appreciate how these rational framing effects can and should be part of good decision-making.

But it is standardly (almost universally, in fact) believed that it is completely irrational to be susceptible to any kind of framing effect. And one of the reasons that frames and framing are held in such low esteem (from the perspective of rationality) is that people have focused primarily on the first group of framing effects – the ones revealed by experimental psychologists and behavioral economists. And there is a very good reason for this focus. The initial experimental work on framing effects was part of a very significant narrative about human irrationality that emerged in the last decades of the twentieth century. Looking at how that narrative emerged gives useful and important background. We turn to it now.

The Litany of Irrationality

Every once in a while, experiments and ideas emerge from a narrow university context and take on a life of their own. One such complex of experiments and ideas has become a powerful narrative in the popular imagination. This narrative emerged originally from experiments on the psychology of reasoning and decision-making and then was subsequently reinforced from areas as apparently divergent as behavioral finance and cognitive neuroscience.

Researchers from these areas and others have converged on the basic idea that human beings are fundamentally flawed reasoners, regularly contravening the basic principles of rationality. Laboratory experiments seem to show that even highly educated and trained individuals regularly
and systematically commit egregious fallacies, flouting fundamental laws of logic and basic principles of probability. Some of the experiments are abstract, but many are not. And expertise seems to be no guarantee of success. Doctors evaluating the probability that patients who test positive for a disease really have that disease seem to fare no better than mathematically sophisticated undergraduates in Ivy League schools doing basic tests of logical competence, or MBA candidates assessing investment strategies.

Some of the leading researchers on human reasoning have made drastic claims (in a typically understated academic style). Richard Nisbett, in one of the earliest salvos in what became known as the rationality wars, said that his and other psychological experiments had “bleak implications for human rationality.”7 The cognitive psychologists Daniel Kahneman (joint winner of the 2002 Nobel Prize in economics) and Amos Tversky summed up their early work on statistical reasoning by saying “for anyone who would wish to view man as a reasonable intuitive statistician, such results are discouraging.”8 Others have been more breathless. The title of journalist David McRaney’s best-selling book You Are Not So Smart speaks for itself. Likewise, Predictably Irrational: The Hidden Forces That Shape Our Decisions, written by the cognitive psychologist and behavioral economist Daniel Ariely.9

This is not just an “academic question.” According to the dominant narrative, poor reasoning and irrational decision-making are particularly acute when it comes to finance and investing. Behavioral economics and behavioral finance are, in essence, academic disciplines founded on the premise that market participants are fundamentally irrational when it comes to spending and investing. This basic premise has become well-established among finance professionals and others who make their living in and around financial markets. The websites of major investment companies such as Vanguard offer introductions to behavioral finance for retail investors and investment professionals.10 The personal finance sections of bookstores and websites are packed with books that offer to help save investors from themselves. A great example (and a very well-written

---

7 Nisbett and Borgida 1975. 8 Kahneman and Tversky 1972.
10 The Vanguard site for financial advisors, for example, contains video tutorials on how investors make decisions and how financial advisors can incorporate “behavioral coaching” into their practice. See the Advisor’s Alpha section of the Vanguard advisors’ website at https://advisors.vanguard.com/VGApp/iip/site/advisor/researchcommentary?page=AdvisorAlpha (accessed 3/28/16).
and thought-provoking book) is *The Little Book of Behavioral Investing: How Not To Be Your Own Worst Enemy* by James Montier.\(^\text{11}\)

This is what I call the *litany of human irrationality*.\(^\text{12}\) The dominant narrative that human reasoning is fundamentally flawed is built on a frequently recited and repeated invocation of experiments and studies. But these experiments and studies are narrowly focused and much more equivocal than generally thought. They have also been over-interpreted. If the case for human irrationality were really as powerful as it has been taken to be, then it would be a miracle that we ever managed to develop financial and economic systems sophisticated enough to allow investors to go astray as spectacularly as they are supposed to do. So I, like quite a few others, think that this is an area where a degree of skepticism is badly needed.\(^\text{13}\)

In any event, while many participants in the “rationality wars” have taken aim at different aspects of the litany of human irrationality, one central part of the litany has been left completely untouched. This is the role of frames and framing in human reasoning, as illustrated in our five examples. We tend to value things as a function of how we frame them. The way in which we look at the world influences how we evaluate our different options and the outcomes that they might bring about. In many cases shifting frames leads us to change how we evaluate things. And this is what leads to framing effects. In a typical framing effect we find ourselves valuing the same thing differently depending upon how we frame it. From the perspective of the psychology of reasoning and behavioral finance (and just about everybody else) susceptibility to framing effects is Exhibit A in the narrative of human irrationality. And even the

---

\(^\text{11}\) Montier 2010.

\(^\text{12}\) I owe a terminological debt to Björn Lomberg, who writes about the environmentalist litany in his book *The Skeptical Environmentalist* (Lomberg 2001).

\(^\text{13}\) Early push-back against the litany came from the philosophers Elliot Sober 1978 and L. Jonathan Cohen 1981, each of whom objected to the basic idea that there could be an experimental demonstration of human irrationality (for example, by arguing that the basic idea of irrationality only makes sense against the background of shared rationality). For further broadsides and commentary from a philosophical perspective see Stich 1990 and Stein 1996 respectively. Objections to the litany have also come from an evolutionary perspective, with authors such as Gigerenzer 1991 arguing that performance on probability tests drastically improves when the tests are presented in terms of frequencies rather than probabilities, which reflects how our brains evolved to deal with probabilistic information. A related objection to the litany comes from the rational analysis approach first developed by the psychologist John Anderson 1990, which starts from the basic premise that the mind is well adapted to its environment. Oaksford and Chater 2007 use rational analysis to explain (away) many of the key data points from the litany.
most dedicated opponents of the litany of irrationality have nothing positive to say about framing effects.

The basic message of this book, however, is that this way of thinking about frames and framing effects is fundamentally mistaken. Framing is completely unavoidable and, while framing effects can sometimes be irrational, they often are not. In fact, framing can be a powerful tool for rational decision-making and problem-solving. Using frames is one of the best tools we have for rational self-control, and it is also key to social cooperation and collaboration.

What Are Frames?

This book is about frames and framing. However, there is no chapter entitled “What are frames?” This is deliberate. The concept of framing is used in many different contexts and in many different disciplines. I am sure that there is no single way of thinking what a frame is that all those who work with the concept would agree upon. The concept of a frame is itself something that can be framed in many different ways, one might say. Trying to give a watertight definition is surely a fool’s errand.

So, for example, we certainly find the concept of a frame in experimental psychology and behavioral economics, as we’ll see in more detail in Chapters 2 and 3. But what psychologists and economists think of as a frame is rather different from the type of framing envisaged by, for example, the cognitive linguist George Lakoff in his best-selling book *Don’t Think of an Elephant: Know Your Values and Frame the Debate*. Economists and experimental psychologists often study how people are influenced by positive versus negative ways of characterizing outcomes. Will you react differently to something that is presented as a direct loss that you might incur (a drop in your wealth from $1,500 to $1,000, for example), as opposed to a gain that you forgo (say, a lost opportunity to increase your wealth from $1,000 to $1,500)? Lakoff, in contrast, is interested in the cognitive metaphors that structure how we think and speak about thorny moral and political issues. His interest is much more in political messaging. In a classic behavioral economics experiment there is a neutral outcome (a level of wealth, for example) that can be characterized in different ways. Not so much for the kind of framing metaphors that Lakoff is interested in. As we will see in the concluding chapters, when issues get sufficiently complicated, there are

---

14 Lakoff 2004.
no neutrally characterizable outcomes. Every description of, say, abortion embodies a particular frame and narrative. Even trying to describe abortion in the way it might be described in a medical textbook will be viewed as embodying a particular stance, which illustrates what I will call the illusion of frame-neutrality.

Many sociologists, political scientists, and communication theorists look at how frames work in communicative contexts – similar in focus to cognitive linguists such as Lakoff, but different in emphasis. They think that linguistic frames are not just metaphors. They function also as mechanisms to perpetuate stereotypes and reinforce power structures, so many social scientists think. The analysis of how news is framed is a key part of media studies and media analysis, itself carried out in different ways by different groups within academia. The issue here is not just how a particular news story is presented (the “slant” that one newspaper or TV station might give to the story), but also how items are selected as newsworthy in the first place.

To add to the confusion, frames are often discussed without being labeled as such. I think of the great German logician and mathematician Gottlob Frege as a pioneer of the idea that how we frame the world determines how we think about it and act within it. One of Frege’s great insights was the notion of sense (Sinn in German), which he also called mode of presentation (Art des Gegebenseins in German, literally meaning way of being given). Here are two typically Fregean examples. It makes a difference whether you think about the number 2 as the smallest prime number or as the square root of 4. It makes a difference whether you think about the planet Venus as the Morning Star or as the Evening Star. Thinkers can take different cognitive perspectives on a single object and those cognitive perspectives can lead them to reason and to act in very different ways. As I bring out in Chapter 5, Frege’s ideas are very important in articulating a framework for thinking about the rationality of framing.

In short, frames are discussed and deployed very differently in psychology, economics, linguistics, sociology, political science, and philosophy. Even a simple catalog of all the different ways that people talk about and appeal to the framing concept would be a book-length study – and one that, as far as I know, nobody has undertaken. It would be futile to try to give a definition that will capture all and only the ways of thinking about frames that would feature in the as yet unwritten Book of Frames. But nor, on the other hand, is it just a linguistic accident that the word “frame” keeps coming up in broadly similar contexts.
Ludwig Wittgenstein, who was an inveterate opponent of philosophers’ attempts to provide necessary and sufficient conditions for key concepts, introduced the idea that certain key concepts might function by family resemblance, rather than having a core meaning that can be stated in a definition. His favorite example was the concept *game*. There is no single feature shared by everything that would be correctly classified as a game, but any game is similar in some key respect to some other games. The unity of the concept *game* is secured by a complex pattern of similarities holding within and across sub-groups of games (multi-player games, solitary games, competitive games, word games, and so on).

I am sure that Wittgenstein is right about games, but not so sure that what he says applies to the concept *frame*. A better model comes, I think, from the fourth century BC philosopher Aristotle, who single-handedly pioneered the systematic study of language and reasoning (among many other things). Aristotle distinguished two categories of words – *synonyms* and *homonyms*. Two words or expressions are synonyms when they have the same meaning, while a word is a homonym if it has multiple unrelated meanings. So, to take examples that have stood the test of time, “bachelor” and “unmarried man” are synonyms, while the word “bank” is a homonym (referring either to a financial institution or to the side of a river, depending upon context). If we think about the multiple ways that people talk about frames, I think it’s clear that we are not dealing with a case of synonymy. The expression “mode of presentation,” as used by Frege, does not have the same meaning as the word “frame,” used by Amos Tversky, for example. But nor is the word “frame” functioning as a homonym when it is used by sociologists as opposed to behavioral economists.

Fortunately, Aristotle can help. He suggested that there is a special category of homonyms. These are homonyms where the multiple meanings share a common core, or what is sometimes called a focal meaning.\(^\text{15}\) There is a common semantic thread, as it were, that runs through the different meanings. This common semantic thread relates them, without collapsing them into one. On Aristotle’s picture, this type of word or concept is unified by more than family resemblances, without having a single meaning that might be captured by a simple definition. I think

\(^{15}\) The expression *focal meaning* is due to Owen 1960, which is an important paper on Aristotle’s theory of meaning. Other Aristotle scholars have acknowledged Owen’s insight, but proposed different terminology. See Shields 1999, for example, who prefers to talk about core-dependent homonyms.
that *frame* and *framing* function in this way. They are used in fundamentally different, non-synonymous ways, but at the heart of those there is a single core.

The best expression of this core that I have encountered comes from an important book published in 1974 by the sociologist Erving Goffman. In *Frame Analysis*, Goffman explores the basic idea that individuals and groups use frames as ways of organizing their experience. He works with a very general definition. For him, frames function as “schemata of interpretation” that allow individuals “to locate, perceive, identify, and label.”16 Each of the ways of thinking about frames discussed above (and all of those that we will be looking at in this book) have this basic characterization at their core – as their focal meaning, in the Aristotelian sense. They can each be viewed as taking this core idea of a schema of interpretation that allows individuals to locate, perceive, identify, and label – and then developing and expanding it in divergent (and often incompatible) ways. Different things can count as schemata of interpretation. They can have different sources. And they can function very differently.

The topic of this book is the rationality of frames. Now that we have said something about frames, it is time to turn to rationality.

**Background: How We Reason vs. How We Ought to Reason**

Continuing with Aristotle, he is often quoted as having said that “man is a rational animal.” In fact, it’s not clear that the phrase “rational animal” actually originated with Aristotle, as opposed to later Stoic philosophers and Aristotle’s own commentators and interpreters in the Middle Ages. But no matter – it’s a useful phrase and a good place to start.

The first thing to say is that talking about rational animals is fundamentally ambiguous. In one sense, when we talk about rationality we are talking about a process – the process of reasoning. In this sense, rational animals are distinctive because they solve problems by reasoning; because they can think about abstract concepts such as justice and truth; and because they are capable of reflecting on themselves and their place in the world. This is a descriptive conception of rationality – rationality as a process. In the descriptive sense, when we say that human beings are rational animals we mean that human beings are distinctive because they engage in certain types of reasoning.

---

In another sense, when we talk about rationality we are talking about an aspiration, an ideal. Rational animals are distinctive in this sense because they understand that thinking and reasoning are subject to standards. Rational thinkers know that they are bound by the laws of logic and other basic principles of sound reasoning, such as the probability calculus. What makes them rational thinkers is that they aspire to abide by those laws and principles. This is a normative conception of rationality. From a normative perspective to say that human beings are rational animals is to say that they are distinctive because they understand (at some level) that reasoning is governed by ideal standards, and they strive to respect those ideal standards.

If we are looking, as the medieval Aristotelians were, for what is distinctive of human beings, then we need to focus on rationality in the normative sense of striving to follow standards rather than the process-driven, descriptive sense. Researchers in comparative psychology and cognitive ethology have identified many types of non-human animal behavior that seem best understood as resulting from mental processes of reasoning. It seems clear that non-human animals are not automata responding blindly and rigidly to external stimuli, although it is not easy to understand the type of reasoning that non-human animals are engaged in.\textsuperscript{17} To the best of my knowledge, though, there is no experimental or field evidence that animals besides language-using humans are rational in the second, normative sense.

In any event, in the human case there is clearly a constant tug-of-war between rationality as process and rationality as ideal. It is not news that we all often reason in ways that fail to respect ideal standards. According to the litany, these are not just occasional misses. The litany is supposed to show that the ideal standards are regularly and spectacularly contravened, so that the normative principles of rationality are more honored in the breach than in the observance. In effect, what the litany says is that the overwhelming evidence from everyday life and scientific studies is that human beings are irrational animals.

The litany is typically embedded in a much richer narrative. According to one version of this richer narrative there are distinctive patterns in our widespread departures from ideal rationality. Manifest irrationality has, as it were, its own rationale. This is where experimental psychology and

\textsuperscript{17} I have tried to clarify some of the issues in Chapters 6 and 7 of my book \textit{Thinking Without Words} (Bermúdez 2003a). See also the essays in the aptly named collection \textit{Rational Animals?} (Hurley and Nudds 2006).
behavioral finance are reinforced by other disciplines. According to influential evolutionary psychologists our departures from the canons of normative rationality have a solid evolutionary explanation. Irrationality is not a random phenomenon. Instead it occurs because what we are actually doing when we reason is applying cognitive short-cuts and rules of thumb that actually helped our ancestors to survive in the physical and social environments they had to deal with. Those cognitive short-cuts and rules of thumb (often called fast and frugal heuristics) sometimes yield the same results as classical normative theories of rationality, but more often they diverge.\(^\text{18}\)

So-called dual-process (or dual-system) theories offer a second way of interpreting the litany.\(^\text{19}\) The key idea here is that problem-solving can engage either of two separate and distinct reasoning systems. Different authors draw the distinction between the two systems in slightly different ways, but the basic contrast is between a fast and intuitive System 1 and a slow and deliberate System 2. Table 1.1 gives some of the characteristics

<table>
<thead>
<tr>
<th>System 1</th>
<th>System 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast</td>
<td>Slow</td>
</tr>
<tr>
<td>Intuitive</td>
<td>Analytic</td>
</tr>
<tr>
<td>Automatic</td>
<td>Controlled</td>
</tr>
<tr>
<td>Concrete</td>
<td>Abstract</td>
</tr>
<tr>
<td>Implicit</td>
<td>Explicit</td>
</tr>
<tr>
<td>Pattern recognition and mental short-cuts</td>
<td>Rules-based</td>
</tr>
<tr>
<td>Shared with animals</td>
<td>Uniquely human</td>
</tr>
<tr>
<td>Reflects specialized knowledge-bases</td>
<td>Linked to general intelligence</td>
</tr>
<tr>
<td>Evolutionarily old</td>
<td>Evolutionarily recent</td>
</tr>
</tbody>
</table>

\(^{18}\) A classic example is the work of the evolutionary psychologists Leda Cosmides and John Tooby. See, for example, Cosmides 1989 and Cosmides and Tooby 1992. For an introduction see §8.3 of Cognitive Science: An Introduction to the Science of the Mind (Bermúdez 2020). For more on fast and frugal heuristics see Gigerenzer, Todd, and Group 1999, Gigerenzer and Selten 2001, and Gigerenzer and Gaissmaier 2011.

\(^{19}\) For influential presentations see for example Stanovich and West 2000 and Kahneman 2011. Evans 2008 defends dual-process theory against some of the most prominent objections.
typically taken to distinguish the two systems and the types of reasoning that each employ.

Both System 1 and System 2 can be mistaken, of course. But, as Table 1.1 shows, only System 2 is supposed to be explicitly guided by normative ideals of rationality. Our characteristic departures from ideal rationality typically occur when System 1 is engaged.

Plainly, therefore, the litany of human irrationality only makes sense against a background of normative theories of ideal reasoning. Such normative theories are developed in many different places. The most prominent are formal logic, statistics, decision theory, and game theory.

**Logicians** study forms of deductive argument, distinguishing those that are valid from those that are invalid. Logically correct reasoning deploys argument forms that preserve truth – where it is impossible for the premises of the argument to be true and the conclusion false. The laws of logic are intended to be completely general, holding whatever the subject matter.

**Statisticians** focus on cases where conclusions are supported by evidence but not entailed by that evidence. They look at different ways in which hypotheses are supported by evidence, and how strong a conclusion it is legitimate to draw from limited evidence. Issues here include how probability is to be understood and estimated, as well as how rational thinkers should update their beliefs when new evidence comes in.

**Decision theorists** explore how rational thinkers ought to choose when confronted with a range of possible actions. Decision-making can take place in many different conditions. Sometimes the outcomes of each possible action are known with certainty. Sometimes decision-makers know only the probabilities of each potential outcome. And sometimes not even the probabilities are known. Different canons of rational choice apply in each case.

**Game theorists** look at the problems and challenges emerging from situations involving collective action – interactions where there are multiple participants and the outcomes are fixed by the actions of all participants. These situations typically cannot be modeled just by decision theory, because the different outcomes depend not just on what each chooser does and how the world turns out, but crucially on how other participants behave and their expectations about other players.

---

20 For informal introductions to the issues logicians grapple with, see Read 1995 and Priest 2000.

21 Statistics is not an area that lends itself to informal introductions, but a good place to start is with the volumes on Statistics and Probability in Oxford University Press’s useful Very Short Introductions series – Hand 2008 and Haigh 2012 respectively.

22 For a short overview see Chapter 1 of Bermúdez 2009. For more detail see Resnik 1987, Binmore 2009a, or Peterson 2009 (which also contains material on game theory).

23 Binmore 2007 is a good place to start. For a more technical and shorter overview see Leyton-Brown and Shoham 2008, which is an abridged version of Shoham and Leyton-Brown 2008.
This book will engage with all of these areas. On the negative side we will look at how framing is supposed to be fundamentally illogical, driving thinkers to make irrational decisions and to weigh evidence incorrectly. As already hinted, these charges need much closer scrutiny. On the positive side we will look at some of the ways that framing can help tackle problems left unsolved by classical normative theories such as decision theory and game theory. So, for example, classical decision theory cannot explain the rationality of self-control. Nor can classical game theory explain why rational agents often choose to cooperate and collaborate in certain basic types of social interaction. We’ll see that bringing frames into the picture yields important insights into how agents actually resist temptation and how they end up cooperating with each other.

But showing that framing effects are useful falls a long way short of showing that they can be *normatively* rational – as opposed to descriptively rational. The controversial claim that I want to make is that being sensitive to how things are framed *can* be perfectly rational *in the normative sense*. I emphasize the word “can” for a reason. According to both proponents and opponents of the litany, framing effects are both widespread and invariably irrational. It would be absurd to respond to this by saying that it is *always* rational to be susceptible to framing effects. What we will be exploring in this book is that there are *some* cases where it is perfectly rational to be sensitive to frames – and indeed, where it might even be irrational *not* to be sensitive to frames. What makes this interesting is that these are very important cases – both practically and theoretically.

Here’s the plan for the book. As I said earlier, one reason why frames and framing are typically looked down upon (from the perspective of rationality) is that people have focused primarily on the first group of framing effects that we looked at right at the beginning of this chapter – the framing effects revealed by experimental psychologists and behavioral economists. We will start out by looking at these.

First, the case for the prosecution. Chapter 2, “Framing: The Classic Experiments,” reviews some of the most influential experiments that have driven discussions of framing and primed the litany of irrationality. Chapter 3, “Where the Rubber Hits the Road: Investors, Frames, and Markets,” extends the discussion to investing behavior and the financial markets, looking at classic examples of the kind of irrational framing that
investment advisors warn against (while at the same time often succumbing to it). And then in Chapter 4, “Juliet’s Principle,” I summarize the case for the irrationality of framing, based on the fundamental claim that it is irrational to allow oneself to evaluate differently two or more different ways of framing the same outcome, when one knows that that is what they are (“a rose by any other name would smell as sweet,” in other words).

With that review in place, the remainder of the book will be spent exploring cases like Agamemnon at Aulis and the self-control example, starting in Chapters 5 and 6. In Chapter 5, “Rational Frames?,” I introduce and discuss in more detail a range of framing effects of the type that I offer as candidates for rational framing effects. Then in Chapter 6, “Agamemnon and Climate Change,” I develop a framework for how those framing effects work and why they can be rational. These two chapters are the theoretical heart of the case for the defense.

With that theoretical framework in place, I turn to more detailed examples and applications. Chapter 7 (“Framing Temptation and Reward: The Challenges of Self-Control”) will take us back to self-control. In Chapters 8 (“Chickens and Chariot Races: Framing in Game Theory”) and 9 (“Fair’s Fair: Framing for Cooperation and Fairness”) we will look at game theory and how frames can work there. We will focus particularly on how games can be framed from the perspective of the individual (the “I”-frame) or from the perspective of the group (the “we”-frame), and how this framing can affect what counts as a rational solution to the game. And then in Chapters 10 and 11 we look at what happens in the sort of discursive deadlock that occurs both privately in personal dilemmas and publicly in debates about controversial topics such as, say, abortion, gun control, and immigration. In Chapter 10 (“Getting Past No: Discursive Deadlock and the Power of Frames”) I diagnose discursive deadlock as rooted in clashes of frames. And then in Chapter 11 (“Opening the Door to Non-Archimedean Reasoning”) I show how a rational, frame-sensitive reasoner can tackle both public and private discursive deadlock.