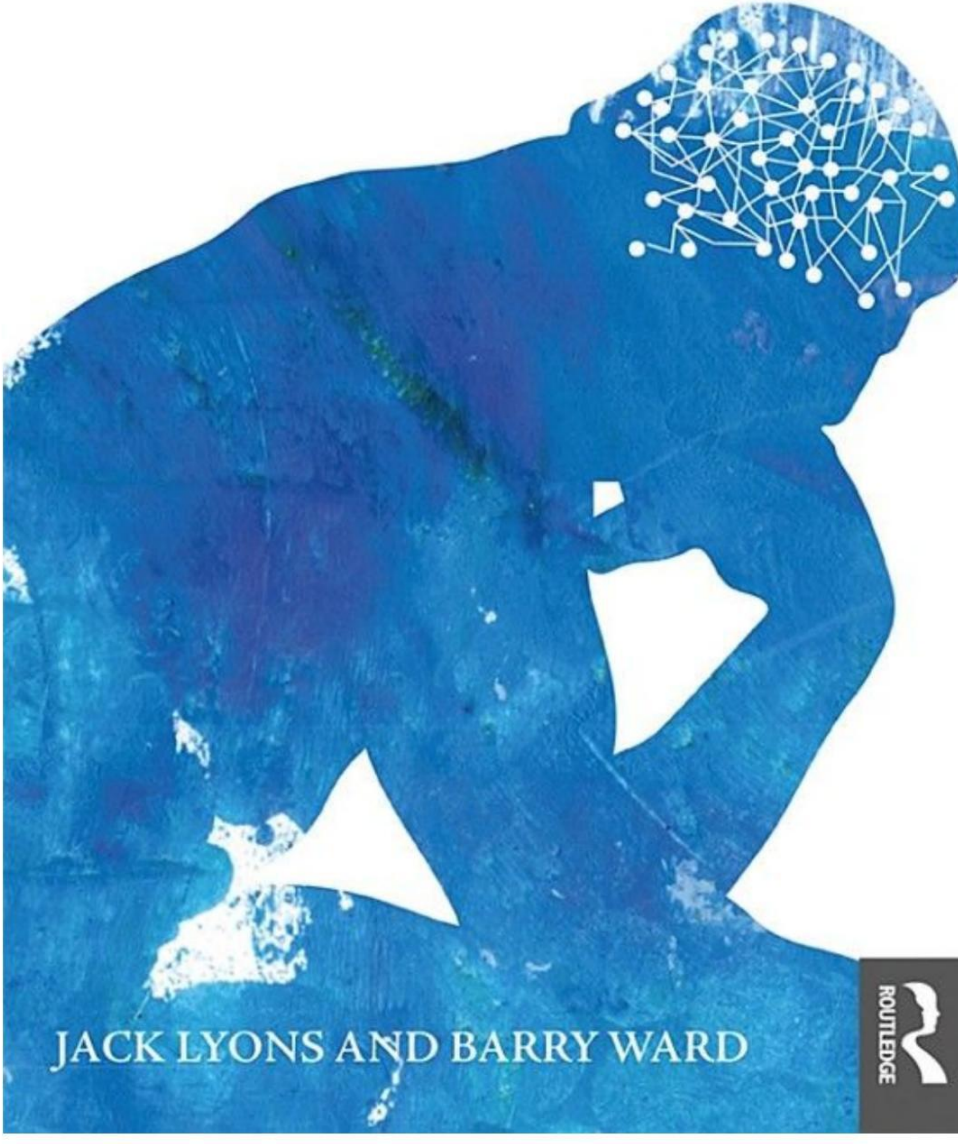
THE NEW CRITICAL THINKING

An Empirically Informed Introduction



JACK LYONS AND BARRY WARD

ROUTLEDGE

First published 2018
by Routledge
711 Third Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication Data
A catalog record for this book has been requested

ISBN: 978-1-138-68747-9 (hbk)
ISBN: 978-1-138-68748-6 (pbk)
ISBN: 978-1-315-54228-7 (ebk)

Typeset in Berling and Futura
by Apex CoVantage, LLC

Visit the companion website: www.routledge.com/cw/Lyons

Contents

Acknowledgments

This book has benefited from conversations and interactions with a number of students, colleagues, and friends. Thanks to Bret Schulte, Stephanie Schulte, Angela Hackstadt, Mike Bishop, Chris Hill, and David Barrett for helpful conversations. Thanks to Ram Neta and some anonymous referees for detailed comments on the manuscript. Thanks to Ted Poston for many helpful comments after having test driven the manuscript with his students at the University of Southern Alabama. Special thanks to Kathryn Zawisza for all kinds of help over the last few years not only with the manuscript but also with the development of a connected curriculum. Thanks also to our department chairs, Tom Senor and Ed Minar, for support and encouragement. Some of the materials here were developed under a Teaching Innovation Grant from the University of Arkansas. For moral support and patience, thanks to Raina Lyons and Traci Parnell-Ward.

Instructor's Preface

This textbook offers a practical, hands-on approach to critical reasoning. Its aim throughout is to make students (mostly non-philosophy majors) better reasoners outside of the classroom. The book is primarily intended for critical thinking courses of the kind that are taught at the freshman and sophomore level at numerous universities and colleges across the U.S. It makes a number of crucial innovations regarding both content and approach. Regarding content, critical thinking texts generally focus almost exclusively on *types of arguments and associated errors of reasoning*. There are two reasons why this approach is inadequate. First, it implicitly assumes that when we form beliefs badly, we do so because we are reasoning badly—we simply lack the knowledge to identify and "weed-out" defective or fallacious arguments. If this diagnosis were correct, the standard approach might be a good one. However, the work of numerous psychologists, notably the Nobel laureates Daniel Kahneman and Amos Tversky, has demonstrated that humans often form mistaken beliefs not because of errors of reasoning, but because of unreliable heuristics/automatic belief-formation processes on which we all, frequently and unreflectively, rely. Thus, standard critical thinking texts are simply blind to a *pervasive* source of intellectual error. By contrast, our approach is psychologically informed both in its diagnosis of the errors' sources and in the relevant remedies. Throughout the text, we focus on the real, psychological roots of erroneous belief. In some cases, merely pointing out the error's source and warning against it is an effective cure. However, in others, psychologists have discovered techniques we can use to

"work around" our natural intellectual blind-spots, and we exploit these techniques where they are most useful, notably regarding probabilistic inference.

Second, the standard approach implicitly assumes that assessing an argument's premises is largely outside the scope of a critical thinking text. This is a major omission. We very often hold others mistaken, not primarily because of errors of reasoning, but because they rely upon sources of information that we find dubious. In this Internet age of clickbait and fake news, everyone has a pulpit, and "journalistic balance" often means that the opinions of those with no relevant expertise are placed on a par with those of evolutionary biologists, climate scientists, and epidemiologists. Thus, the standard critical thinking textbook advice to trust the opinions of relevant experts when there is a consensus among them, doesn't begin to address the relevant issues. Many students are either incapable of making any evaluation of who *the real* experts are, or will have made their own idiosyncratic selections of experts on highly dubious grounds. We take the issue of expertise, and of how to deal with Internet sources in particular, seriously. So, we include detailed and upto-date discussions of the nature of science, and more fundamentally, general discussion of belief on the basis of testimony and how suitably structured communities of experts can legitimately earn the trust of reasonable people. Thus, the book has an epistemological component that is generally absent from other critical thinking texts and is indispensable for real-world argument evaluation.

In addition to these fundamental differences in content and approach, there are substantial innovations in our presentation and development of the standard elements of a critical thinking text, most notably regarding argument evaluation and identification. Textbooks typically, indeed almost universally, present identification before evaluation. However, in this book, the order is reversed. This is not a bug; it's a feature. The identification of the various components of an argument in the most humdrum text or speech generally hinges on the judgment that if these components

formed an argument, it would be a passably sensible one: argument identification generally presupposes argument evaluation. Even in the most *simple*, *realistic* examples, a passage will often contain, in addition to some genuine argument, a wealth of thematically related material that contributes nothing to the argument's strength, and hence, is irrelevant. It is almost perverse to ask students to winnow out the argument without significant prior instruction on evaluation. In particular, the idea that indicator words provide remotely adequate guidance is a pedagogical fiction that, in our experience, largely hinders and confuses students who are genuinely seeking to understand and develop the relevant skill. Hence, our approach.

We also provide an original (at least as far as critical thinking texts are concerned) and intuitively compelling way of proving the validity of a broad class of valid arguments and argument forms: explicitly demonstrate that it is impossible to make the conclusion false while making all the premises true. Crucially, this technique does not rely upon truth tables, thus avoiding the damaging confusion that often arises in critical thinking classes when English indicative conditionals are, counterintuitively, analyzed as material conditionals i.e., using the material conditional truth table. Further, it provides a unified, comprehensive treatment of both simple truth-functional argument forms and categorical syllogisms, eliminating the unnecessary complication of analyzing the latter using Venn diagrams. Many instructors and *some* students love the formal simplicity of truth tables and Venn diagrams, but many students don't. Thus, the main text, instead, employs the above technique and also simple Euler diagrams. For instructors who prefer truth tables and/or Venn diagrams, we cover the same material in those terms in appendices A and B respectively.

We have also sought to address a number of difficulties that will be familiar to anyone who has spent some time in the trenches, teaching this demanding, and yet, elementary course.

1 Getting students to understand and successfully apply the core concept of deductive logic, validity, is not easy. Numerous instructors will report that, *at the end of their course*, a high percentage of their students still don't get it. We have located two problems here: properly understanding the notion of *possibility* invoked in the definition of validity and recognizing that the *actual* truth values of the premises and conclusion of an argument are not a guide to its validity. We target both issues with careful discussion and vivid, worked examples.

2 Relating the theory of causal inference to real-world cases can be challenging. Our response here is two-pronged. One strand integrates causal inference with a discussion of counterfactual conditionals. This is not the norm in critical thinking texts, but it renders the treatment more intuitive, since counterfactuals are ubiquitous in causal inference. The other strand emphasizes the statistical nature of much causal inference, focusing on the difficulties of inferring causation from correlation, and the importance of confounds. This gives the student the basic conceptual resources to evaluate the kinds of scientific studies that are frequently ignored in critical thinking texts.

3 While a basic grasp of probabilistic and statistical concepts is invaluable for understanding the world we live in, the mathematical aspect of these subjects intimidates many students. Also, psychologists have established that our intuitive probabilistic inference is highly error prone. In response, we use frequency trees to provide a user-friendly way of assessing common probabilistic information, illustrating the take-home lessons of Bayes's theorem, notably the base rate fallacy, while employing only elementary arithmetic. For students who are comfortable with more math, we give an elementary formal discussion of the probability calculus and Bayes's

theorem.

4 Real-world argumentation is typically more a matter of back-and-forth than of static lists of numbered premises followed by conclusions, and a traditional approach to critical thinking can be hard to relate to real-world debate. By introducing some of the epistemology of defeasible reasoning, we teach students which sorts of dialectical moves are legitimate, which are not, and more generally, which approaches can direct discussion towards finding the truth.

This book (and the related web-based materials) includes more content than most instructors would consider covering in a typical one-semester course. This allows the instructor to customize things to some degree. The core of the book, as we see it, consists of section 2 of the Introduction—where the dual systems psychology is introduced, Chapters 1–4 and Chapter 7, which cover foundational material on deductive and inductive arguments, and argument identification and reconstruction. (Parts of Chapter 7 will be *a little* puzzling without having read Chapters 5 and 6.)

A course tailored to science majors should certainly include Chapters 5, 6, and 9, which respectively cover causal inference, probability and statistics, and some elementary philosophy of science. One aimed at humanities students should probably include Chapters 8, 10, and 11, which respectively deal with testimony and the proper evaluation of source material, rhetoric, and dialectic.

Someone who wanted to entirely omit the more difficult and formal material could teach the Introduction, Chapter 1, Chapters 4–5, and 7–11. Chapter 1 provides an adequate basic account of deductive inference, and the later chapters cover standard inductive inferences, causal inference, testimony, and the evaluation of sources, science, rhetoric, and dialectic, while omitting the more sophisticated treatment of deduction and the treatment of probability altogether.

Introduction to Critical Thinking

If you've ever tried to convince someone of something, or if anyone has ever tried to convince you of something, then this book is for you.

This book is concerned with critical reasoning and the proper evaluation of reasoning, argument, and persuasion by others. Our focus is more on the consumer than the producer; that is, our goal is not so much to make you better at convincing others but to make you better at being convinced only when you ought to be convinced.

1. The Aims and Causes of Belief

To convince someone is to influence their beliefs. Why would anyone want to do that? For one thing, it's sometimes nice to have people agree with you; that's some reason to convince them to see things your way. More importantly, however, our *behavior* is guided by our beliefs. The actions we take are more a matter of what we believe than of how the world really is. Beliefs are things that can be true or false. When they're false, there's a mismatch between the belief and the world. In those cases, it's the belief that determines behavior, not the world itself. If I can convince you to *believe* that my car is worth $2,000, then you'll be inclined to pay that much for it—*regardless of how much the car is really worth.* If you believe that politician x is the best candidate for the job, then you'll be inclined to vote for x, *whether x really is the best*

candidate or not. And so on.

Our concern in this book is with belief and the proper regulation of belief. To start out, it is worthwhile to consider a few general questions about the nature of belief:

- what is belief (i.e., what is it to believe something)?
- what is the aim of belief?
- what are the causes of belief?
- what are the *proper* causes of belief?
- to what extent are our beliefs under our voluntary control?

We'll offer answers to these questions, but since this is a critical thinking book, after all, we don't expect you to just accept what we say.

First, there are a number of different ways we might try to define belief, but the simplest is probably this: to believe something is to think it is true. That may not seem immensely helpful, but we want to make explicit the relation of belief and truth.

This becomes especially relevant when we turn to our second question: what is the aim of belief? Here we need to tread carefully. There may be more than one aim of belief, especially depending on the answer to the last question. If we have a fair amount of control over our beliefs, we might adopt various beliefs for any number of different reasons. I might adopt a certain religious belief because I want to fit in with my family (or to minimize my chances of going to hell!); I might choose to believe that I'm going to win the race in order to increase my chances of doing so; morality might require me to have certain charitable beliefs about other people.

Still, we think that there is one *primary* aim of belief, even if there are a number of secondary aims. That is truth; *the primary aim of belief is truth.* To put it differently, the purpose of belief is to get at the truth; beliefs are *supposed to be* true.

We can say this in a somewhat more nuanced way. ***Epistemology*** is the study of knowledge and rational belief, and ***"epistemic"*** means "having to do with knowledge." Some think we sometimes have good moral reasons to believe certain things—i.e., it would be immoral to believe otherwise. It is even more plausible that we can sometimes have good practical, or pragmatic, reasons to believe certain things—e.g., if I believe my product is superior, I will be a more effective salesperson, and that is in my interest for a number of reasons, etc. And sometimes what you ought—morally—to believe or ought—pragmatically—to believe can conflict with what you ought—*epistemically*—to believe. A better way to put the claim that the primary aim of belief is truth is this: *the* ***epistemic*** *aim of belief is truth.* What you ought—epistemically—to believe is what is likely to be true.

Our concerns in this book are epistemological, so we will ignore moral and pragmatic considerations regarding belief. This is not, of course, to say that they're not important, only that they're not part of our current concern. There are very difficult questions about how to resolve conflicting commands. What, for instance, should we do when epistemic norms and pragmatic norms push us in different directions? Suppose it's in my practical and professional interest to believe that my client is innocent, but all the evidence indicates that he's not. What should I believe? Pragmatically, I should believe he's innocent; epistemically, I should believe he's not. OK, but what—all things considered, and without the qualifications—should I believe? We don't think there's an easy answer to this question. Although we are about to leave behind the topics of moral and pragmatic norms for beliefs, focusing entirely on epistemic norms, this doesn't mean that we think that epistemic norms always override other norms. We think it's very important to do our best to believe only what is true, but we don't think it's the only thing in life that's important. We are trying to offer a guidebook to thinking, not a guidebook to living.

So what are the causes of belief? Why do people believe the things they do? There are a lot of answers here: because they see

it; because they remember it; because someone told them so; because it follows logically from other things they believe, and so on. Lumping things together, we can say that one cause of belief is *evidence*. As we will understand it, *the **evidence** for some belief is any consideration that makes that belief more likely to be true.* If evidence were the only cause of belief, our task here would be somewhat easier. We will have a lot to say about what counts as evidence (that is, good evidence), and this is all of what most critical thinking books try to do. But since people believe things from other causes as well, we should address some of these other causes.

One obvious cause of belief is training, or *upbringing*. Surely you've noticed that people tend to believe what their parents believed; this is especially apparent in the case of political and religious beliefs. It is unlikely that this is entirely a matter of being exposed to a particular body of evidence. Children are taught certain things before they are old enough to question them, and these beliefs become deeply ingrained and difficult to question, let alone eradicate.

We seem to also be sensitive to mere *repetition*; if a politician or advertising agency wants us to believe something, they will repeat the message again and again, because they know it works. If we hear something enough, we are likely to start to believe it. Obviously, however, this doesn't provide us with *evidence*, with any real reason to think the claim is true. Repetition doesn't make the claim any more likely to be true; it just makes us more likely to think that it's true.

There are a number of *emotional factors* that incline us to believe certain things.

- If I really like you, I'm more likely to believe what you tell me than if I don't.
- If you're saying something that I want to be true, I'm more likely to believe it than if I don't want it to be true. This, of course, is known as wishful thinking.

- We have all known people who are too trusting of their spouse or significant other. Sometimes this is the result of wishful thinking (they don't want the partner to be unfaithful so convince themselves that the partner isn't). But sometimes it is the result of a different emotion: loyalty. If I feel guilty about *not* believing something, this may make me more likely to believe it.
- People derive some of their sense of identity from their membership in certain groups or subcultures. It is important to one's self-image to think of oneself as Catholic, or on the debate team, or a Cubs fan, or a Democrat, or what have you. Surely our membership in these groups makes us more likely to believe like the rest of the group.

In extreme cases, disease or *brain damage* can cause beliefs. A certain kind of chemical imbalance can cause people to think that others are out to get them or that someone else is telepathically inserting thoughts into one's head.

You can probably think of a number of others. And if you've studied psychology, you'll know of many more. Some of these we will discuss later. For the present purposes, a few examples are sufficient. We have the following, partial, list:

- evidence
- upbringing
- repetition
- emotion
- brain damage

In contrast with the first item, the others have one glaring feature in common: *they are unreliable*. Suppose you hold the political views that you do entirely because your parents raised you to believe as they did. Clearly, had you been raised by different parents, you'd have very different political views. You might as

well decide your political views by rolling dice. Obviously, unless you have some independent reason to think your parents are better at these things than other parents, believing something just because they did is a terrible way of getting at the truth. Similarly with the other causes: saying something over and over doesn't make it more likely to be true; our emotional reactions are determined by many factors other than the truth, and brain damage doesn't ordinarily make someone *better* at getting things right.

Often, these non-evidential factors operate in tandem with other causes, including evidence. Perhaps not many people believe things *entirely* because their parents did. The point, however, is that insofar as these non-evidential causes are exerting an influence, it's a bad influence, since they don't do anything to increase the likelihood that the belief is true.

Recall that the (epistemic) aim of belief is truth. Insofar as these non-evidential causes of belief are not truth-conducive, we should not be influenced by them. *Of all the causes of belief, the only (epistemically) proper cause of belief is evidence.* Because our beliefs are often caused or at least influenced by illicit factors, we need to be wary about why we hold a given belief.

Here's where things get tricky.

It is often said that nearly everyone dies of the same thing: namely, a lack of oxygen to the brain. The idea here is that even though there are many different *distal* causes of death, the *proximate* cause is always the same. Drowning (the distal cause) causes death by keeping air out of the lungs, so the blood runs out of oxygen, thereby depriving the brain of oxygen. Heart attacks do it by stopping the heart, so the blood doesn't pump, so the brain is deprived of oxygen. Strokes block oxygenated blood from getting to the brain; severe blood loss results in an inadequate supply of oxygenated blood to pump to the brain, etc.

Similarly, people believe nearly all of what they do believe for the same reason: namely, it seems true to them. That is,

something's seeming true is nearly always the proximate cause of belief. There might be some people who are capable of just choosing to voluntarily believe certain things—and we mean genuinely believe, not just say they do, or have it be their official position, or act as if they believed. But such cases are probably very rare. When we engage in wishful thinking, we are generally unaware that we are doing so. We are influenced by our desire without realizing that the desire is influencing us. I don't *decide* to believe that my chances of winning the lottery are good; I just find myself believing it, or at least being inclined to believe it. My chances of winning just *seem* good. Similarly, my in-group members have certain beliefs about gun control, and those views *seem sensible* to me; their positions *seem plausible*; their claims *seem correct;* they have the "ring of truth," even if there's no real evidence in their favor.

The problem is that this "truthiness" (to co-opt a term from Stephen Colbert), this feature whereby certain things just seem true to us, accompanies pretty much all belief (probably all involuntary belief). Whether the belief results from evidence, emotion, brain damage, or the rest, it will seem true. So, the mere fact that something seems true is, by itself, no reason at all to think it really is true. *Truthiness is not evidence.* As we have seen, things seem true to us for all sorts of reasons that don't have any bearing on truth, so the mere fact that something seems true should not lead us to think that it actually is true. As the bumper sticker says, "Don't believe everything you think."

2. Reasoning and Dual Systems Theory

For the last few decades, and especially in recent years, there has been a convergence of evidence from various independent lines of research into the human mind and brain, that the mind can be

divided into two systems: an ***intuitive processor*** and a ***conscious rule interpreter***. These are more commonly referred to as ***System 1*** and ***System 2***, respectively.[1]

System 1 (the intuitive processor)—so called, even though it's probably a host of distinct processors rather than a single one—operates quickly, automatically, and unconsciously. The kinds of problems that it solves are generally either evolutionarily old problems or newer problems that the individual has a great deal of practice with. The answers come immediately and without deliberation. In math class, by contrast, you were told to "show your work," to leave a record of the processes or steps by which you came to a particular result. System 1 doesn't "show its work"; it delivers an answer only, without you being able to tell how it came to that result. That's why it's called "intuitive": an intuition is a sense that something is true, where the person can't tell you why he or she has that sense. The person "just knows."

System 2 (the conscious rule interpreter), on the other hand, is slow, deliberate, effortful, and able to solve a wide variety of new and unanticipated problems. System 2 does show its work; when you are using it, you know what steps you are going through to solve the problem at hand; you know what your reasoning is.

The terms "intuitive processor (/System 1)" and "conscious rule interpreter (/System 2)" refer to the *sources* of beliefs. It is useful to have terms for the *process* of coming to believe by way of these sources, and for this, we will use the terms "***intuition***" and "***reasoning***," respectively. Keep in mind that from here on out, these will be technical terms with these highly specific meanings.

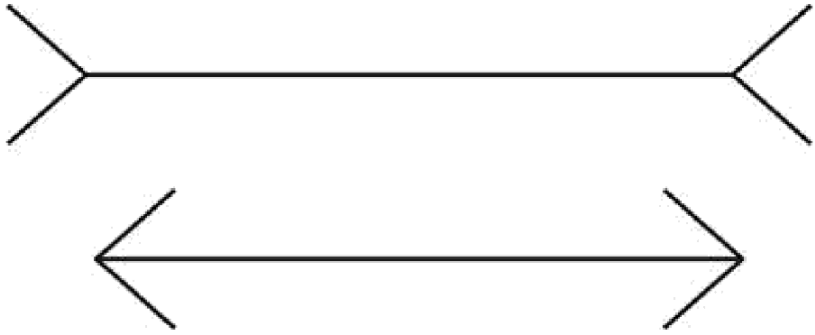
Compare playing tennis to playing chess. When you're playing tennis, you're relying largely on intuition. You concentrate on, say, hitting the ball, but you don't concentrate on how to go about hitting the ball. To hit the ball, you have to move your arm, and to do that, you have to contract certain muscles while relaxing other muscles, but this isn't something you think about consciously; this is done on autopilot. In fact, if you try to think about it, it will make you worse rather than better. (If you want an unfair

advantage in tennis, or pool, or something similar, try to get your opponent to think about the things she normally does on autopilot: conscious deliberation here just makes everything worse.) Tennis, of course, is *physically* effortful, and certain aspects of it, like tracking the ball, are *cognitively* effortful; but certain aspects of it, like getting your limbs into the right places, are cognitively effortless.

Chess, on the other hand, is not something most of us can do on autopilot. We have to painstakingly reason out the possible consequences of each move, thinking about what the opponent might do in response. We know and consciously apply the rules of the game (bishops can move only along diagonals, etc.); we have to reason it out. We are aware of the mental effort that goes into making a decision to move a piece, and unless we forget, can usually tell others why we chose to make the move we made. Some chess experts, apparently, have learned the game so well that they have trained their intuitive processors to "just see" things that the rest of us would have to laboriously reason out. For most of us, however, chess is a good illustration of an activity that requires effort and attention.

Reasoning in general (i.e., employment of the conscious rule interpreter) is perceptibly *effortful*; we can't just start the process and let it go, but we must keep providing impetus at every step. Intuition (i.e., use of System 1), on the other hand, is effortless; we normally don't even have to get it started—it goes off by itself—and we certainly don't have to nudge it along.

Intuition is like perception in a number of important ways; they are both fast, automatic, evolutionarily old, and perform complex processing without our having any access to their inner workings. We often even describe intuition in perceptual terms: "I could just *see* that he was going to cheat on her." Another similarity is that both perception and intuition offer up appearances, or seemings (sometimes also called "impressions"), as their final outputs. Consider a case of perceptual illusion. When we look at the following familiar illusion



we have a strong sense that the top line is longer than the bottom line. It *seems* longer; it *looks* longer. We can determine, by measuring, that the lines are actually the same length, but even then, the top line continues to look longer. Even though we don't *believe* that the top line is longer, it still *seems* to be longer. Similarly, we need not accept, or believe, the outputs of intuition. The fact that something seems intuitively to be true doesn't mean that we're required to believe it.

The tennis example might make it seem as if System 1 is concerned largely with motor skills, but our main concern is its role in delivering intuitive judgments about the world. Consider the following question:

> If a baseball and bat together cost $1.10, and the bat costs a dollar more than the ball, how much does the ball cost?

If you are like most people (including most Princeton students), you answered 10¢ and did so pretty quickly. But now reason it out; do the math. If the ball is 10¢ and the bat is a dollar more than that, then the bat must be… and so the combination must be…. Now we see that the right answer has to be that the ball costs 5¢; this makes the bat cost $1.05, for the grand total of $1.10. If you answered 10¢ the first time, that is because this is the answer System 1 delivered. It was quick, effortless, and the answer seemed right. But then we engaged our System 2 and reasoned out the correct answer. Here, then, is an illustration of the two systems in action.

In this particular case, intuition gets it wrong, and reasoning gets it right. In other cases, the reverse is true. For certain kinds of tasks, intuition is actually quite reliable (see Malcolm Gladwell's *Blink*), but for other tasks, it's pretty unreliable (see Kahneman's *Thinking, Fast and Slow*).[2] It is an interesting question whether, on balance, reasoning is better at getting things right than is intuition. But we won't try to answer that question, because for our purposes, it doesn't matter. You want not only to be likely to get things right, but *to have some assurance* that you're likely to get things right. For this, you need to bring in System 2.

The end-product of System 1 is an appearance, or a seeming: a certain claim seems true or seems plausible, at least. But as we noticed before, "the" intuitive processor is actually a collection of separate systems, some of which may be reliable in a given instance and some of which may not. By its very nature, intuition is silent as to how the beliefs—or appearances—it produces are generated. If something just seems true to me, then I can be fairly sure this is the result of intuition. But some intuitions are reliable and some are not, and I can't tell which this one is, because I generally don't know anything about the intuition except that I have it. I can't tell, for instance, whether or not it is the result of wishful thinking or a learned expertise, precisely because System 1 doesn't show its work. Intuition and perception involve unconscious operations, but because the operations are unconscious, we don't have access to the grounds on which the system delivers its verdict. There are good (reliable) intuitions and bad (unreliable) intuitions, but you can't tell the difference just from the nature of the intuition. They *all* seem right.

Reasoning is not *necessarily* any more reliable than intuition. Whether your conscious rule interpreter is reliable or not depends almost entirely on what rules you're following. The difference between reasoning and intuition, however, is that when reasoning, you are in a position to know which rules you are following, and so, if you know which rules are the right ones, you can be sure you're doing it right. Furthermore, you can learn the rules simply

by reading, understanding, and remembering. This makes System 2 trainable in a way that System 1 is not. The purpose of this book is to teach you which rules are the right ones, thereby helping you to train and improve your System 2. It is likely that certain aspects of reasoning will thereby eventually become so deeply ingrained that, like the chessmaster, you'll start to "just see" the strengths or flaws in certain inferences, strengths or flaws that used to require careful reflection and consideration to spot. If so, you'll be training System 1 *by means of* training System 2. In any case, it's System 2 that is our focus, because that's the one that can be trained by the articulation of rules.

We'll say more as we go on about the two systems, especially about how the inner workings of System 1 can lead you astray and how the proper use of System 2 can play a corrective role. This is an area of active empirical research, and there is lively disagreement among the researchers at the frontiers about a number of important details. We will restrict ourselves to those claims for which there is broad consensus in the field. The version of "dual systems theory" (also known as "dual process theory") that is presented here, and which will emerge in more detail in later chapters, is a summary view, simplified for our pedagogical purposes.[3] The basic insight, drawn for diverse corners of psychology and neuroscience, is that for many cognitive tasks, we have both a fast, intuitive, unconscious method of solving the problem, and a slow, effortful, rule-based method. This much of the theory, at least, is pretty uncontroversial, and we think it is here to stay. We have noted that thinking of "the" intuitive processor as a single entity is probably an oversimplification. For our purposes, it doesn't really matter. In fact, you're welcome to think of System 1 as a little person sitting on your shoulder and whispering answers to you if it helps you to remember this central fact: the problems that we solve unconsciously and intuitively are often solved unreliably (in systematic and predictable ways to be discussed in later chapters), and we want to be critical and reflective about these solutions. We don't want to just accept them

at face value. We don't want to just believe whatever System 1 "tells" us.

What this warning amounts to is this: the mere fact that something seems true to you is very little reason to think it either is true or is even likely to be true. Things seem true to us for all sorts of reasons, some good and some bad. If you want to get at the truth, be leery of claims that seem right even if you can't say why. Beware the "ring of truth."

3. Reasoning, Evidence, and Arguments

Reasoning, as we have defined it, is the conscious, deliberate articulation and assessment of evidence. It's a mental process, a thing that happens inside your head. An ***argument*** is a set of statements consisting of one or more *premises* and a *conclusion*. The ***premises*** are statements that are offered as *evidence* for the conclusion, and the ***conclusion*** is the statement whose truth the argument is intended to establish. Arguments are typically offered by someone in an effort to convince someone else of the conclusion. Arguments and reasoning are the two main things this book will focus on: the presentation of evidence (arguments) and the evaluation of the evidence thus presented (reasoning). Reasoning is, of course, also involved in the *construction* of arguments, which is a topic we will treat indirectly. Much of what's involved in constructing good arguments is critically evaluating the arguments that occur to you and in filling in the gaps in those arguments. These are topics we will address directly and at great length.

We will, of course, have much to say about arguments as the book unfolds, but let's warm up to things a bit by looking at some simple examples. Suppose you and I are discussing our plans for tonight and I say the following:

> I bet Sarah won't come out with us tonight. Her grandfather is in town visiting her, and she probably won't want to either bring him out on the town or ditch him to hang out with us.

This is an argument. I'm trying to convince you of a conclusion by presenting evidence for it, in the form of premises. More specifically, the conclusion is that Sarah won't come out with us tonight. (The "I bet" functions to indicate that I'm drawing an inference and also to hedge my bets a little, to express a bit of uncertainty.) The premises are that her grandfather is in town, and that she won't want to bring him or ditch him. To make things explicit, we will write this out as follows:

> (P1) Sarah's grandfather is in town.
> (P2) Sarah won't want to leave her grandfather alone, and she won't want to bring him out.
> (C) Therefore, Sarah won't come out with us tonight.

As you can see, we've numbered the premises, so that we can refer back to them. And we've explicitly marked the conclusion, by labeling it (C), but also by including the word "therefore," which wasn't in the original passage but very clearly indicates that a conclusion is being drawn.

There's often more than one right way to do this. We could have, for instance, broken up (P2) into two separate premises:

> (P2) Sarah won't want to leave her grandfather alone.
> (P3) She won't want to bring him out.

Either way of doing it is fine.

Contrast this argument with the following:

> I'm not going out with you guys tonight. My grandfather is in town visiting me, and I don't want to either bring him out on the town or ditch him to hang out with you.

Though similar in obvious ways, this one isn't an argument. I'm

not trying to convince you of anything; in particular, I'm not trying to provide *evidence* for the claim that I'm not going out with you tonight. My telling you that should be good enough (if not, the rest of what I say would hardly convince you, because you probably won't take my word for that either). Rather, the stuff about my grandfather is intended to *explain* why I'm not going out, not to provide *evidence* for it.

What's the difference? In the argument, the fact that her grandfather is in town is offered as a reason to believe that Sarah won't join us tonight; it's a reason to think that (C) is true. As for the premises, I'm simply *telling* you that Sarah's grandfather is in town. I'm laying that down as a starting point. I'm trusting you to take my word for it, and if you don't, then you're not going to believe much of the rest of what I say either. That's part of what's meant by calling (P1) a *premise*: it's a *starting point*, something we have to agree to, or at least assume for the sake of argument, if the rest of what I say is going to be productive. In contrast, I'm not asking you to take my word for the claim that Sarah isn't going to join us tonight; I'm using (P1) and the other premise(s) to provide *evidence* for this claim, which is why it's labeled as a conclusion. In this way, I'm offering *support* for the conclusion: if you trust me on the premises, you don't have to trust me on the conclusion; the premises make it reasonable to believe the conclusion.

Like many of the examples we'll use in this book, this example is artificial (although recognizably the sort of thing you might encounter in real life). Because it's artificial, there's no surrounding context. We could easily imagine this same passage about Sarah in a context where you already knew (P1), (P2), and (P3). In that case, I'm not *informing* you of (P1) etc.; I'm *reminding* you. But they're still premises. They're still starting points on which to base the conclusion.

Contrast all this with the second passage. The claims are all thematically related, but none of them are obviously serving as evidence for the others. The claim that I'm not going out tonight is

something I'm telling you but not really arguing for. I'm trusting you to take my word for it. I'm not drawing an inference; I'm stating an intention. I'm not hazarding a prediction; I'm declaring my decision. My statement that I'm not going out tonight is more like a premise than like a conclusion, in that it's something I'm simply telling you. However, we wouldn't treat it as a premise, because it isn't offered as evidence for anything.

Exercises 0.3

For the following passages, say whether they embody arguments or not. If so, say which statements are serving as premises and which as conclusions, and explain how the premises provide evidence for the conclusion, i.e., why the premises would give someone reason to think the conclusion is true.

1 There's a pangolin. I *love* pangolins!

2 That's a pangolin. Only pangolins have scales like that.

3 I can't believe my brother voted for that fascist in the last election! I don't respect him anymore, and we're certainly not going to talk politics at Thanksgiving, if I can help it.

4 You shouldn't eat that.

5 Some mushrooms that look like that are deadly poisonous. You shouldn't eat that, unless you're sure what it is.

6 If you keep driving like this, you're going to get into an accident. And anyway, it's very hard on the car. You should slow down and quit driving like such a jerk!

7 We stayed at the party until 4 A.M. And I drank way too much. That's why I slept in until noon.

8 Ugh, you can't root for Arkansas! I'm an LSU fan!

4. Why Reason (Properly)?

Most critical thinking books focus on teaching students how to reason properly. This is certainly important, and most of the pages here will be aimed explicitly at this. But in a way, this focus omits the most important part of reasoning: the readiness to do it at all. You can't do it right if you don't do it at all, and your natural inclination—*our* natural inclination—is not to reason at all. (Recall that "reasoning" here is a technical term; in some sense, of course, you're reasoning nearly all the time.) Reasoning, as we have seen, takes effort, while intuition doesn't. Intuition will supply you with an answer, and it will feel like the right answer (it seems true), whether it is or not. Although we are going to spend most of the rest of this book trying to explain how to reason well, it is likely that some of the gains you experience will come about simply as the result of your trying to reason well, remembering to engage System 2. Much of the difficulty people have with reasoning is not so much not knowing *how* to do it, but not thinking to do it at all. You already know enough about proper reasoning that if you even just try to reason (as opposed to merely forming beliefs) you will do much better than if you don't—whether you learn anything from the rest of this book or not. Of course, that's no reason to refuse to read on. A semester's practice will get you in the habit of engaging System 2, and we do think we can teach you things that will make that system more efficient and more accurate.

"But why should I want to reason properly, or to reason at all, for that matter?" This is a very good question. Because reasoning is an effortful activity, we automatically have a certain inclination to avoid it. Yet some things that are effortful are pleasant anyway.

If this weren't true, few people would engage in physical exercise or puzzle solving. We hope that reasoning will become more pleasant as this course goes on.

But even when reasoning is not enjoyable, it's a good thing to do. This is because proper reasoning makes you more likely to achieve the truth than improper reasoning or not reasoning at all. "Why should I care about the truth?" you might ask. Again, good question. There's a two-part reply. The first part is that much of the time, you simply *do* care about the truth. If you want to know whether smoking causes cancer, you want to know *the truth* about whether smoking causes cancer. You don't just want the first answer that pops into your mind; you want the *right* answer. Most questions are, if they are worth answering, worth answering correctly, i.e., worth answering with the truth.

The second part of the reply is that, whether you care about the truth or not, you *should* care about the truth. You have various practical concerns, and having true beliefs about the world makes it much easier to meet those concerns than having false beliefs. Sure, you'll sometimes get lucky, and there are odd cases where someone gets what they want because they were wrong about something. But all in all, being right about things will get you more of what you want than being wrong about things. If you're interested in the relation between smoking and cancer, that may be because you're trying to decide whether to smoke or not (or to quit smoking or not). Quitting smoking is an awful, miserable thing to do, but it's not nearly as awful or miserable as having lung cancer. So, if smoking really causes cancer, then I want to quit. If it doesn't, then I don't. Again, I don't want to just have a belief about the relation between smoking and cancer; I want (or should want) the truth.

There is a final reason to want to develop the practice of engaging System 2. System 1 behaves in certain predictable ways, and this makes it possible for unscrupulous people to co-opt that system, more or less *stealing* your assent without first paying an appropriate price. Because System 1 is automatic, people can

generate intuitions in you without your consent, and if you're not on guard against it, they can thereby make you do things that you don't really want to do. Clever people know how to elicit intuitions in you, in a way that bypasses, or that fails to engage, System 2. We will examine this in more detail later, but for now, consider advertisers, who know that if they can create a mental association in you between their product and, say, puppies, you'll be more likely to buy their product. So, they bombard you with images of their product's name alongside cute puppies, and when you're in the store and you see that product, it creates a warm and fuzzy feeling in you that makes you want to buy that product. There is nothing stopping the purveyors of an inferior product from running puppy ads, so this ploy can cause you to buy an inferior product. But you didn't *really* want to buy the inferior product, so they got you to do something you didn't want to do!

We sometimes have the mental image of a bull with a ring through its nose, which ring is presumably for the purpose of making an otherwise formidable beast into a docile animal that can be easily led wherever one may want it to go. The ring is great for the rancher, but is bad for the bull; he would be better off without it, unless the rancher happens to be nice. Having a System 1 that is primed and raring to go, while your System 2 sits idle and unused, is like having a ring in your nose, one that allows other people to lead you wherever they see fit. This might work out well, provided that the clever people around you have your best interests in mind. But they often don't. Even if they did, most of us would want to be the masters of our own fates. Our freedom and autonomy depend on our ability to engage our reasoning processes and to not rely entirely on intuition. Insofar as we uncritically believe whatever intuition tells us, our beliefs—and therefore our actions—are determined by whatever forces happen to surround us, be they benevolent, selfish, or uncaring. It is only if we engage our reasoning processes (which will rarely, if ever, be automatically activated by the world around us) that we take control of the direction and character of our believing and acting.

It is only in this way that we really become full-fledged agents, deliberately acting in the world, rather than being acted on by it and reacting reflexively.

5. Plan for the Book

Here's where we're headed: the cleanest, most straightforward, and strongest type of argument is a ***deductive*** argument: it aims for absolute certainty, and the evidence it provides, if successful, is conclusive and final. Part I (Chapters 1–3) deals with deduction; we spell out this notion of absolute certainty in more detail and provide an introduction to *logic*, in mostly non-formal terms. We examine some practical methods to determine whether or not a deductive argument is successful. We draw special attention to some of the more common deductive argument forms, so you'll recognize them when you encounter them later. Part I ends with a discussion of how to identify and reconstruct deductive arguments from real cases in print, speech, etc.

Deductive arguments are not the only form of argument there is, and they're not even the most common. An ***inductive*** argument is one that doesn't attempt to provide absolute certainty and finality but only reasonable probability. In some cases, this probability can be *quite* high, and there's nothing wrong with an inductive argument, simply because it isn't deductive. The strength of an inductive argument will come in degrees, however, and some do little or nothing to establish their conclusions. Part II (Chapters 4–7) is concerned with inductive arguments. We discuss the standard forms of inductive argument, explaining the standards for properly evaluating them. We give special attention to reasoning about causes and to reasoning that is explicitly about probabilities and/or frequencies. As with Part I, Part II ends with a treatment of the identification and reconstruction of real-life arguments and the proper use of the information learned here

about induction.

Many critical thinking texts restrict their subject matter to reasoning, leaving students on their own (or in someone else's care) for picking the right premises or starting points. We think this is a mistake; we think it is possible to give some general instruction on choosing the right starting assumptions for arguments. In Part III (Chapters 8 and 9) we discuss the nature and credibility of science and the more general issue of believing something on the basis of someone else's vouching for its truth. Most of us aren't in a position to directly evaluate much of the evidence on which most scientific claims are held, but if we understand better the nature of the scientific enterprise, we will know better why—and when—to put our trust in science.

Part IV focuses on argumentation: the real-life interactions of people trying to convince each other (or trying to remain unconvinced). Chapter 10 is on ***rhetoric***: the art or practice of persuasion, often by means other than cogent argumentation. We approach rhetoric from the consumer's angle, rather than the producer's: we aren't trying to make you better at persuading others through non-rational means; our aim is to help keep you from being taken in by it. Chapter 11 deals with ***dialectic***: the back-and-forth of argumentative "moves," with all the dynamic messiness that it entails—a messiness overlooked by a tendency to view arguments simply as collections of sentences.

Finally, the book closes with a short glossary of the reasoning fallacies: mistakes common enough that they have names. The book is accompanied by some online appendices, where some of the issues that we gloss over in the main text receive more detailed and advanced treatment, and where solutions to some of the exercises are given.

Summary

Beliefs, though sometimes false, are essentially connected to truth in at least the following sense: to believe something is to think it is true. Belief goes wrong when it turns out false; the epistemic aim of belief is truth. ("***Epistemic***" means "having to do with knowledge.") Though belief has many causes, the only epistemically proper cause is evidence. ***Evidence*** is any consideration that makes that belief more likely to be true. Because claims can seem to be true for many sorts of reasons besides evidence (e.g., wishful thinking, habit, partisanship), the mere fact that something seems true is, by itself, no reason at all to think it really is true.

Well-established research in psychology and the cognitive sciences divides the mind into an ***intuitive processor*** (aka ***System 1***), which generates ***intuitions***, and a ***conscious rule interpreter*** (aka ***System 2***), the use of which we will call ***reasoning***. The former is fast, automatic, and unconscious in the sense that it doesn't "show its work." Because of this, we can't tell which of our intuitions are reliable and which are not. Because we want to know this, we have to rely on System 2 as much as possible. System 2, however, is only reliable if it's equipped with, and properly using, the right rules. The main subject of this book is these rules and their proper use.

An ***argument*** is a set of statements consisting of one or more premises and a conclusion. The ***premises*** are statements that are offered as evidence for the conclusion, and the ***conclusion*** is the statement whose truth the argument is intended to establish.

Why should you care about reasoning properly, or reasoning at all? One answer is that it's a pleasurable, though challenging, activity. Another is that it makes you much more likely to get at the truth about things, and having true beliefs will result in you getting much more of the other things you want in life than having false beliefs would. Most importantly, perhaps, by developing the habit of engaging System 2, you make it harder for others to manipulate you; you thus become a more fully autonomous agent, rather than a pawn of someone else.

Notes

1 The terms ‘intuitive processor’ and ‘conscious rule interpreter’ are from Smolensky, Paul (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11 (1):1–23. Kahneman, Daniel. *Thinking, fast and slow*. Macmillan, 2011; Kahneman, Daniel, and Frederick, Shane (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psycho logy of intuitive judgment, 49*, 49–81 use ‘System 1’ and ‘System 2’. This latter terminology is fairly unmemorable, but it might help to keep in mind that the intuitive system comes *before*—both in evolutionary history and the development from childhood to adulthood—the conscious one. Some researchers use different terminology.

2 Gladwell, Malcolm. *Blink: The power of thinking without thinking*. Back Bay Books, 2007. Kahneman, Daniel. *Thinking, fast and slow*. Macmillan, 2011.

3 To read more about dual systems theory, see Kahneman, Daniel, and Frederick, Shane, op cit.

Part I
Deduction

Chapter 1
Validity and Why It Matters

The main topic of this book is the proper evaluation of evidence. This means the proper evaluation of arguments.

An ***argument***, again, is a set of sentences (or, as we'll sometimes call them: claims, statements, or propositions) consisting of one or more *premises* and a *conclusion*. The ***premises*** are statements that are offered as *evidence* for the conclusion, and the ***conclusion*** is the statement whose truth the argument is intended to establish. Logicians typically distinguish between deductive arguments and inductive arguments. Roughly speaking, an argument is ***deductive*** if the truth of premises would guarantee the truth of the conclusion; an argument is ***inductive*** if the truth of premises would render the truth of the conclusion probable, without guaranteeing it. Some inductive arguments are very powerful, and the probability they confer is extremely high. There's nothing *wrong* with an inductive argument just because it doesn't absolutely guarantee its conclusion. Nevertheless, inductive arguments are messier and more complicated than deductive arguments. Thus, in this chapter and Chapters 2 and 3, we will focus on the stronger and simpler kind of argument, the **deductive** argument. Simple doesn't mean easy, and the next two chapters will be a bit abstract, but please bear with us. The skills and concepts mastered here will be important for nearly all other reasoning.

Our provisional understanding of deduction is rough in two ways. First, we'll want to say quite a lot more about what's meant

by "guarantee." Second, if we were to define deductive and inductive arguments as those that guarantee or make probable their conclusions, it would follow that there couldn't be *bad* arguments of either type, arguments that abjectly fail to provide the kind of support they're intended to. Consequently, we'll officially define deduction and induction in terms of the *aims* of the argument, that is, in terms of the *intentions* of the person offering the argument. Thus, we will define a ***deductive argument*** as one that *aims* at *validity*, i.e., one that purports to be valid. "Validity," of course, is a technical term that replaces the more intuitive but less precise "guarantee." Just what it means is the topic of this chapter. (We'll say more about aims and intentions a bit in this chapter, but more so in Chapter 3.)

1. Distinguishing the Good from the Bad

The goal here is to distinguish good arguments, ones whose premises provide a genuine reason to believe the conclusion, from bad ones. The good news: You already know a lot about how to do this. From an early age, we reliably use this ability on a daily basis. So, the task of this book is not to introduce some alien, intellectual discipline, but to develop and refine a skill you already possess. To see that we have this skill, take the following pair of examples:

(P1) All members of species X have lungs.
(P2) y is a member of species X.
(C) Therefore, y has lungs.

(P1) All members of species X have lungs.
(P2) y has lungs.
(C) Therefore, y is a member of species X.

The first is a good argument and the second is a bad one, and we

confidently make that judgment. So, in some sense, we already know the difference.

The question now is: can we say what the difference is? What is it about the good argument that makes it good and the bad one that makes it bad? What is the *relevant contrast* between them, the difference that makes a difference?

What counts as a successful answer here? First, we want to know what makes arguments good *in general*, not just the first argument in particular. It is relevant and true to say, "the first argument is good, because "y has lungs" follows from y being a type X and all Xs having lungs," but that answer is too specific. It does not tell us how to evaluate arguments about economics or physics or the likelihood of rain. There's another problem with that answer. To say that the conclusion *follows from* the premises is correct, but unhelpful. If we can't say what that means in simpler terms, saying "it follows" is no more illuminating than saying the argument is good. We haven't *explained* what it is for the argument to be good. The same goes for saying that the conclusion *is a consequence of* the premises, or that the premises *imply* the conclusion, or, if you've already been exposed to some logic, that the argument is *valid*. All true, but they won't explain the idea to someone who genuinely lacks the ability to discriminate the good from the bad, or help us better understand the nature of good argumentation so we can improve our own ability.

We can sneak up on the problem by focusing on the bad argument. It has a hole in it: it could be that all Xs have lungs, but there are other species that also have lungs, and so, y could be one of those. If Xs are dogs, and cats also have lungs, then maybe y is a cat. So, the conclusion would be false.

As it happens, lots of species have lungs. But even if there weren't any other species with lungs, the premises leave open the possibility that such species exist and that is enough for the argument to have a hole in it. And this hole is what makes it a bad

argument. The first argument is good because it has no hole; it's airtight: if all type Xs have lungs, and y is an X, the conclusion that y has lungs is inescapable.

There's something to this, but unfortunately, talk of holes is just a metaphor here, and so, it's too wooly to provide precise guidance. There's not *literally* a hole in the argument, as when we say there's a hole in the wall or in my sweater. For someone who doesn't already have the skill of evaluating arguments, telling them to look for holes is vague, hand-waving advice.

But it does capture something important. So, we need to figure out the precise idea to which the metaphor points. Here's a way of putting it: what is special about the first argument is that the truth of the premises would *absolutely guarantee* the truth of the conclusion; if the premises were true, the conclusion would have to be true. Or, to put it most precisely: it is *impossible* for both the premises to be *true* and the conclusion to be *false* together. This statement is non-metaphorical, and it explains the goodness of the argument in simple terms that do not presuppose specialized logical knowledge: *impossible, true*, and *false*. We call arguments like this "valid."

An argument is ***valid*** if and only if it is impossible for the premises to be true and the conclusion to be false together.

The other ones, the ones that lack this special property, we call "***invalid***."

This definition fits with our two examples. What makes the first argument good is that it is absolutely impossible for it to be false that y has lungs, given that it is true that y is an X and all Xs have lungs. What makes the second example bad is that it clearly is possible for all Xs to have lungs, and for y to have lungs, and yet for y not to be an X (i.e., for it to be false that y is an X). And if we fail to recognize the disconnect between the second argument's premises and its conclusion, we are clearly allowing ourselves to be misled, to be persuaded by premises that just don't provide a good reason to believe the conclusion. On the other hand, if we allow ourselves to believe the conclusion of the first argument,

given its premises, we make no such error.

More generally, in life we typically want to believe truths *and only* truths. To have any success at that goal, we need to have some kind of policy for deciding what to believe. Here's one policy: every time you are confronted with a proposition, flip a coin. If the coin comes up heads, undertake to believe the proposition; if it comes up tails, don't. This is an obviously bad policy. If you followed it, any truths you came to believe would be a matter of sheer luck, and if you acted on the beliefs you acquired, you probably wouldn't do very well. "Eating the rat poison will be a nutritious and delicious experience": Let's flip a coin.

We need a policy that tracks the truth: picks out truths and avoids falsehoods. Picking out valid arguments and rejecting invalid ones is part of such a policy, a crucial component of it. However, just paying attention to validity is not enough. Validity on its own provides *no reason* to believe the conclusion is true. And this is made explicit in our definition: all it says is that a valid argument can't have *true premises* and a false conclusion. It guarantees *conditional support* between the premises and the conclusion: *If* the premises are true, *then* the conclusion must be too. *If not*, all bets are off. Valid arguments with false conclusions are not hard to find. For instance:

> (P1) All human beings have tentacles.
> (P2) All creatures with tentacles live in the sea.
> (C) So, all human beings live in the sea.

It's valid, but it provides no reason to believe the conclusion. Why? Because one of the premises is obviously false, and valid reasoning from a false premise provides no reason whatsoever to believe that we have a true conclusion.

Another way of putting this is to say that valid arguments are ***truth-preserving***: all true premises guarantee a true conclusion: Truth in; truth out. Falsehood in; who knows? (Unless you're

lucky, a false conclusion.) Certainly, the argument gives you no reason to believe it true. So, our policy for truth-tracking should be this: believe only the conclusions of arguments that are valid *and* that have all true premises. These arguments are important enough that we need a name for them. We'll say an argument is *sound* if and only if it is valid and has all true premises. A sound argument must have a true conclusion: *Truth-preservation + all true premises* guarantees a true conclusion.

A word of caution: these two features of an argument (i) its validity, and (ii) the actual truth values of its premises, *have nothing to do with each other*. It is worth emphasizing this point, as people often mistakenly think that the actual *truth values* of the premises and conclusion—whether the premises and conclusion happen to be true or false—can tell us whether the argument is valid or not. But this is not so. For example:

> (P1) Beethoven's music is excellent.
> (P2) If someone's music is still well-known centuries after their death, their music must be excellent.
> (C) So, Beethoven's music is still well-known centuries after his death.

It's perfectly possible that someone might have written excellent music and also that music only survives the test to time if it is truly excellent, and yet some great composer could be unlucky enough for their work to be lost or destroyed before achieving any popularity, and so never be well-known. That's not how it was for Beethoven, but it could have happened. So, the argument is invalid, and yet the premises and conclusion are all plausibly true. Validity and true premises guarantee a true conclusion, but it doesn't work the other way around: true premises and a true conclusion guarantee nothing about the quality of reasoning.

Just to hammer home the point, let's return to our first pair of arguments:

(P1) All members of species X have lungs.
(P2) y is a member of species X.
(C) Therefore, y has lungs.

(P1) All members of species X have lungs.
(P2) y has lungs.
(C) Therefore, y is a member of species X.

You confidently judged the first valid and the second invalid. But notice, there's no way for you to even assign truth values to the premises and conclusions, *because we never even said what X and y are*. If X is dogs and y is Lassie then the sentences in both arguments are all true. If X is monarch butterflies and y is Charlie the tuna, then they're all false. None of this changes the fact that both arguments of the first kind are valid and that both arguments of the second kind are invalid. The actual truth values are *irrelevant* to assessing validity.

To summarize: how well we are reasoning from a set of assumptions does not depend on whether or not they happen to be true. When you assess validity, you should *completely ignore* whether the premises and conclusion actually happen to be true or false. Consider only the information specified in the premises and in the conclusion and determine whether there is any way at all that things could be as stated in the premises and not as stated in the conclusion. If so, it's invalid; otherwise, it is not.

Exercises 1.1

A. **Evaluate whether the following are valid or invalid:**

1 Americans landed on the moon in 1969. No Russians landed on the moon before 1969. So, Americans were the first to

land on the moon.

2 It is always cloudy when it rains. It is cloudy now. So, it is raining now.

3 Americans first landed on the moon in 1972. No one landed on the moon before them. So, Americans were the first to land on the moon.

4 Mary is Pat's sister. So, Pat is Mary's brother.

5 All fish live in the sea. All things that have scales live in the sea. Therefore, all fish have scales.

6 No carnivore is an herbivore. John is a carnivore. Therefore, John is not an herbivore.

7 John is a friend of Brian. Brian is a friend of Jim. So, John is a friend of Jim.

8 Mary is Pat's sister. So, Pat is Mary's sister.

9 Mary is Pat's sister. So, Pat is Mary's sibling.

10 Obama is not the current president. If George W. Bush was the last president, then Obama is the current president. So, George W. Bush was not the last president.

11 If Nadal wins in straight sets in the final, he will win the tournament. Nadal will not win in straight sets in the final. So, Nadal will not win the tournament.

12 Dogs are bigger than cats. Therefore, cats are smaller than dogs.

B. **Evaluate each of the following arguments for validity *and* soundness:**

1 The Eiffel tower is in Paris. Paris is in France. So, the Eiffel tower is in France.

2 The Eiffel tower is in Berlin. Berlin is in France. So, the Eiffel tower is in France.

3 The Eiffel tower is in France. Paris is in France. So, the Eiffel tower is in Paris.

4 The Eiffel tower is in Berlin. Berlin is in Germany. So, the Eiffel tower is in Germany.

5 All birds have wings. All things that can fly have wings. So, all birds can fly.

6 Some dogs are pets. Some pets have four legs and a tail. So, dogs have four legs and a tail.

7 All snakes are poisonous. Pythons are snakes. So, pythons are poisonous.

8 Copper is a metal. All electrical conductors are metals. So, copper is an electrical conductor.

9 All metals are electrical conductors. Copper is a metal. So, copper is an electrical conductor.

10 India is the most populous country in Asia. China is the most populous country in the world. The world includes Asia. So, China is not in Asia.

11 Either Bernie or Hillary is going to get the nomination. Bernie can't beat Hillary. Therefore, Hillary is going to get the nomination.

12 It costs about $.50/mile to drive the average car, figuring in gas, maintenance, depreciation, and the like. Therefore, if you drive the average car and work 5 days a week, 50 weeks a year, moving 10 miles closer to work will save you about $2500/year, keeping everything else the

same.

13 Obama says he's in favor of gun safety legislation. Anyone who says that means to take our guns away. So, Obama wants to take our guns away.

14 If the mind is entirely physical, it ought to be possible to create artificial minds in computers. It is possible to create artificial minds in computers. Therefore, the mind is entirely physical.

C. **More on validity and soundness.**

1 Consider the following argument:

> If John has pancreatic cancer, then he will be dead within 12 months.
> John has pancreatic cancer.
> So, John will be dead within 12 months.

(a) Is it valid or invalid?

(b) Suppose a new cure has been found for pancreatic cancer that is 100% effective and available to all patients. Given that information, what should we plausibly say about the above argument? Is it valid or invalid? Is it sound or unsound? Explain your answer.

(c) *Don't* suppose that a cure has been found. Nevertheless, suppose John survives for more than 12 months, i.e., it turns out that the conclusion is false. What should we now say about the argument? Is it valid or invalid? Is it

sound or unsound? Explain your answer.

2 Which of the following are possible? If impossible explain why. If a genuine possibility, provide an argument as an illustrative example.

1 An argument that is sound and invalid.
2 An argument that is valid and has a false conclusion.
3 An invalid argument with a true conclusion.
4 An argument that is unsound and valid.
5 An invalid argument with true premises and a true conclusion
6 An argument that is sound and has a false conclusion.
7 An argument with a true conclusion that is unsound.
8 An argument with true premises and a true conclusion that is unsound.
9 A valid argument with a true conclusion and at least one false premise.
10 An argument that has false premises and a true conclusion that is invalid.

2. Validity and Impossibility

An argument is valid if, and only if, it is *impossible* for its premises to be true and its conclusion to be false. We favored this definition, because it explains a sophisticated idea, validity, in simpler terms: *impossible, true, false*. However, the first of these terms is not so simple, and we need to talk about exactly what it means.

Here's the concern. In assessing the current state of the U.S. military, a general might say, "*It is not possible* for the U.S. to successfully fight two full-scale wars at once." However, if asked whether with additional investment the U.S. could fight two full-scale wars, the very same general might say, "Sure, *it's possible* for the U.S. to successfully fight two full-scale wars at once." Superficially, it might look like she's contradicting herself—she's asserted the very same thing to be both possible and impossible—but there's no contradiction here. In the first case, the general is taking one set of background information—the current resources of the U.S. army—as a given, and that puts certain restrictions on what is possible. In the second case, the general is not taking that as background, but considering what is possible if those resources were expanded through investment, and that gives a different specification of what is possible. If we spell out what is meant by the two sentences, they come out as "*given the U.S.'s current military capability*, this is not possible" and "*given more investment in the military*, this is possible." The italicized bits are not explicitly stated by the general but are intended to be understood by whomever she is talking to, presumably on the basis of whatever has already been said in the conversation. So, when we use the word "possible" we almost invariably mean: possible *given the background information assumed in this context*. Different background assumptions; different meaning. The word "possible" is ambiguous: in different contexts, *it means different things*.

Some notions of possibility are sufficiently important to have their own names. It is ***physically impossible*** to travel faster than the speed of light. When we say this we are claiming, rightly or wrongly, that going faster than light is incompatible with the laws of physics. It is ***psychologically impossible*** for mice to do algebra. When we say this, we are not claiming that mice have merely lacked the incentive or education to solve simultaneous equations: we're saying that such an ability is simply incompatible with the mental equipment possessed by mice. It is ***physiologically impossible*** for humans to breathe underwater (without special equipment), meaning that breathing underwater is incompatible with the capabilities of the human body. When we use these names for particular types of possibility, we are signaling which background information should be assumed in understanding our use of "impossible": physically impossible, one should assume the (known) laws of physics; physiologically impossible, assume the capabilities of the human body, and so on. Maybe the most common of these specialized notions of possibility is epistemic possibility (recall that epistemology is the study of knowledge and rational belief). To say that something is ***epistemically possible*** (for me) is to say that it's compatible with everything I know. Since you and I know different things, what's epistemically possible for me need not be epistemically possible for you, and vice versa.

Which notion of possibility is the right one for the definition of validity? Referring back to the start of the chapter, our first example is a valid argument and the second is invalid, and that's just a fact. We feel no temptation to say that the second is invalid in one context and valid in another. So, it's not going to be some squishy, context-dependent notion of possibility, where whether or not an argument is valid depends upon whatever background information *happens to be* assumed in a given context. Instead, it will be a special notion of possibility, one that assumes *no background information whatsoever*. Let's call it "***logical possibility***."

Validity is the gold standard of reasoning. When we say an argument is valid, we are saying that you can rely upon that line of reasoning *in any context whatsoever*. We are saying that it is impossible for the premises to be true and the conclusion to be false, *without assuming any unstated background information*. So, logical impossibility is not incompatibility with the facts about physics or psychology or any other body of background information. What logical impossibility boils down to is literal inconceivability. It means that we literally *cannot make sense of* such a possibility. It may be physically impossible for things to go faster than the speed of light, but it's not logically impossible; if the laws of physics were different, say they were the ones we thought were true before relativity theory, then things could go faster than light. It is indeed physiologically impossible for humans to breathe underwater, but it's entirely conceivable that humans could have gills or some other organs that would allow us to do so. So, it's logically possible. When you want to assess what is logically possible, bracket all your background knowledge about how things actually are and consider whether you can coherently conceive of the possibility in question.

Here's a helpful, intuitive way of thinking about it. To say that something is logically possible is to say that an omnipotent being could make it true. This includes all kinds of absurd situations: it's logically possible for me to walk on water, for my car to sprout wings and fly, for cats to talk, and so on. But it doesn't include *everything* as possible. Even an omnipotent god couldn't make a four-sided triangle, a chair that is partly green but also completely red, a man that is taller than himself, etc. There's no contradiction in the idea of an omnipotent god that can't make me taller than myself. There is just *no sense to be made* of a man who is taller than himself. So, there just is no such possibility that our hypothetically all-powerful being is incapable of making happen. So, there's no conflict with omnipotence.

Returning to our original example,

(P1) All members of species X have lungs.
(P2) y is a member of species X.
(C) Therefore, y has lungs.

What this argument has going for it, which you instinctively realized made it a good argument, is that we can make no sense of the premises both being true and the conclusion being false. *Without assuming any unstated background information whatsoever*, it is impossible for the premises to be true and the conclusion to be false. It's *logically* impossible, and that's what we mean by validity.

By adopting a very inclusive conception of possibility, we end up with a very exclusive conception of impossibility. This ensures that validity (because it's defined in terms of this conception of impossibility) embodies the highest level of praise possible; *there simply couldn't be a kind of conditional support better than validity*. That is why we say validity is the gold standard of reasoning. Nevertheless, you might think it ridiculous to demand such a high standard, to define validity using a notion of possibility on which we allow that pigs could fly and cats could talk. The following arguments are *invalid* precisely because such phenomena are logically possible:

(P1) Slypork is a pig.
(C) So, Slypork can't fly.

(P1) Jeoffrey is a cat.
(C) So, Jeoffrey can't talk.

Yes, both lines of reasoning seem very sensible, *given our common-sense background knowledge about pigs and cats*: if you tell me only that Slypork is a pig, I will readily accept on that basis that he can't fly. But that's beside the point here. We've already explained one reason why: a valid argument is one whose reasoning *can be relied upon in any situation whatsoever*, because it does not

depend upon any background assumptions. These arguments simply don't have that general reliability. Why it matters, why we are insisting on validity—at least for now—is something we'll explain further in section 1.5. For the moment, let's just accept the gold standard.

To summarize the last two sections: assessing validity is an exercise in abstraction in two different ways. First, because validity is a matter of *conditional* support, we need to forget about whether the premises or conclusion are actually true. Second, because validity is a matter of logical possibility, we need to forget a host of background information about the world: what the laws of physics and psychology are, whether humans can breathe underwater, etc. We just focus on what the premises and conclusion say, and determine whether it is in any way conceivable for the premises to be true and the conclusion to be false together.

Now that you understand the technical concept of validity, there's just one more small point to understanding our definition of a ***deductive*** argument as one that aims at validity, i.e., that purports to be valid. Obviously, *arguments* don't have goals or aims, but the people who offer them do. We'll see in more detail in Chapter 3 that understanding arguments requires understanding the intentions of the authors of those arguments. For now, it's enough to note that we'll count an argument as deductive if the author *meant* for it to be valid. Thus, there will be *failed* deductive arguments: arguments that aimed at validity but aren't valid.

Box 1.1 Possibility and Necessity

Possibility and necessity are interdefined. Something is possible if and only if it's not necessarily false. Something is necessary if and only if it's not possibly false. Something is ***contingent*** if and only if it's neither necessarily true nor necessarily false.

Box 1.2 Some Important Notions of Possibility

Logically possible: conceivable without contradiction
Physically possible: compatible with the laws of physics
***Epistemically possible* (for S)**: compatible with everything S knows

Exercises 1.2

A. **For each of the following, specify whether the kind of possibility/impossibility at issue is logical or not. If not, say what we might call it, using our examples as guides (so, although we didn't mention "biological possibility," surely there is such a thing as compatibility with the laws of biology, and so on). If it doesn't easily admit of a name, indicate what the relevant background assumptions are.**

1. It's impossible for a fire to burn without oxygen.
2. Dogs can't face their "palms" toward each other.
3. A triangle must have three sides.
4. You can't go to jail for cheating on your spouse.
5. Sound must travel slower than light.
6. If Margaret and Joe are both here, then Joe must be here.
7. You can't survive a zombie bite without

becoming a zombie yourself.

B. **For the following pairs of sentences, determine whether it is logically possible that both be true at the same time. If so, offer a scenario that explains how. Note, this will sometimes require outlandish, though not self-contradictory, suppositions.**

1. It's raining. The streets are dry.
2. I dropped this rock. It never hit the ground or any other surface.
3. Jessica died this morning. She was a half-hour late to work this afternoon.
4. Women are smarter than men. Men score higher on IQ tests than women.
5. Everyone loves Taylor Swift and her music. Taylor Swift is terribly unpopular.
6. Sam Shepard is the greatest English-speaking playwright of all time. Shakespeare existed.

C. **The following pairs couldn't possibly both be true at the same time. In each case, how would you explain this fact to someone who didn't yet see why?**

1. X is a triangle. X has four sides.
2. Sandy and Jules came to the party. Sandy didn't come to the party.
3. If there's an open flame, there's oxygen. There's an open flame but no oxygen.
4. I am taller than Lewis. I *am* Lewis.
5. Rene thinks he might be dreaming. Rene doesn't exist.
6. Things are going to get better, or they're

going to get worse. Things are going to stay exactly the same.

7. I'm in Paris, and Paris is in France. I'm not in France.
8. Nobody's ever run a four-minute mile. Jolene ran a mile in 3:54.

D. **Evaluate the validity of the following arguments:**

1. A cat falls from the top of a tall building and strikes the ground at over 100 mph. Therefore, it is seriously injured or killed.
2. The Empire State Building is made entirely of soap bubbles. All soap bubbles disintegrate within 5 minutes. No new soap bubbles are made. So, there will be no Empire State Building in 5 minutes.
3. John, a human being, is entirely submerged in water for 10 hours without any kind of breathing apparatus. So, John dies.
4. All human beings live in pineapples under the sea. Everyone that lives in a pineapple under the sea is a friend of SpongeBob. So, every human being is a friend of SpongeBob.
5. John is unmarried. So, John is a bachelor.

3. More on Logical Impossibility

We have defined a valid argument as one for which it is *logically* impossible for the premises to be true and the conclusion to be false, and that means that the truth of the premises guarantee the truth of the conclusion, *without our assuming any background information whatsoever*. However, there's a way in which this requirement might seem confusing. Take the following argument:

(P1) John is a bachelor.
(C) Therefore, John is unmarried.

If anything is a valid argument, this is. However, someone might be concerned that we used background information in judging it as valid. To make that judgment, we needed to know that a bachelor is someone who is unmarried—we needed to *assume background information* about what "bachelor" means. This is true, but it's unavoidable and shouldn't be troubling: if we don't allow background information about what the words in the argument mean, *we can't even understand what the argument is saying*, let alone evaluate it. So, being a little more careful, what we mean is that when you evaluate the possibility of the premises being true and the conclusion being false, the *only* background information allowed is the information required for understanding the sentences that make up the argument. Consider the following four arguments:

(P1) The earth's orbit is a circle.
(C) So, the earth's orbit has no corners.

(P1) My father is 40 years older than I am.
(C) So, I am 40 years younger than my father.

(P1) Ice is just solid water.
(C) So, ice will melt when heated above 32° F.

(P1) The cue-ball struck the 8-ball with great force.
(C) So, the 8-ball moved.

Valid or invalid? The first two are valid because of what the terms in the arguments mean. A circle is a curve on which all points are equidistant from one point, the center, and a curve like that is smooth; it has no corners. And if you don't know that, either you don't understand what "circle" means or you don't understand what "corner" means. Given that you do understand, the conclusion is inescapable, Similarly, if my father is 40 years older than me, that just means that I am 40 years younger than him. Once you understand what is meant by "older than" and "younger than," it's inconceivable that the premise is true and that the conclusion is false.

Let's look at the second pair of arguments. We all know that water is liquid if it is above its freezing point, 32° F, and hence, given that ice is solid water, heating it will yield liquid water, i.e., the ice will melt. That's a piece of knowledge about the world that we have had for so long that we have likely forgotten when first we learned it, but that's not the same thing as being part of the meaning of "water." We can see this pretty easily. Suppose we found that some chemical process resulted in a new kind of ice, a solid form of water with a different crystal structure to regular ice. And suppose that if you heat crystals of that ice, they don't liquefy until you heat them above 110° F. We would not describe that situation by saying that the new ice crystals aren't really water. We would say, "Here's a fascinating surprise about water: you can make water ice that doesn't melt even when you push its temperature way above 32° F." In fact, this is the possibility considered in Kurt Vonnegut's novel *Cat's Cradle*. A scientist discovers a novel crystal structure for water, called Ice-9, which is solid at high temperatures. (Spoiler alert: a crystal of Ice-9 is dropped into the oceans, solidifying all of earth's water, triggering an ecological disaster, and the end of humanity.)

The second invalid argument is very similar. We all know that if one pool ball hits another with great force in normal circumstances, it will cause it to move. But again, the given premise does not render the conclusion's falsehood impossible in

the relevant sense; it's the given premise, plus that additional background information that guarantees the conclusion's truth. We can easily form new valid arguments by incorporating these background beliefs as premises, giving:

(P1) Ice is just solid water.
(P2) Solid water will always melt when heated above 32° F.
(C) So, ice will melt when heated above 32° F.

(P1) The cue-ball struck the 8-ball with great force.
(P2) Given the circumstances, if the 8-ball was struck with great force, it moved.
(C) So, the 8-ball moved.

These are now valid, but the originals were not.

Note an important difference between the last two invalid arguments and the two valid ones that preceded them: we *could* give those the same treatment. For example:

(P1) The earth's orbit is a circle.
(P2) Circles don't have corners.
(C) So, the earth's orbit has no corners.

This makes the original reasoning a bit more explicit, but note: it doesn't improve the conditional support for the conclusion. It *couldn't*, because the original argument is valid, and validity can't be improved on. If you don't know anything about the world, but I tell you that the earth's orbit is a circle, then I've already told you all you need to validly infer that the earth's orbit has no corners. I don't need to tell you that circles don't have corners, because if you didn't already know that, then *you didn't understand the rest of the argument.* You can't understand what a circle is and what a corner is unless you know that circles have no corners. Thus, P2 is dispensable, because it doesn't—couldn't!—tell you anything new.

By contrast, in the ice and 8-ball cases, the (P2)s are not logically necessary truths, and they're not things you must already

know in order to understand the (P1)s. If you don't know anything about the world, but I tell you that ice is just solid water, I haven't told you everything you need to know to infer that ice will melt at 32° F. *You'd still need to be told that water melts at 32° F.* You could understand the terms "water," "ice," "melt," and "32" without knowing this.

Sometimes it's hard to know exactly where to draw the line between contingent, factual background knowledge of the sort that we're not allowed to presuppose in assessing validity, and necessary, meaning-related knowledge that just makes explicit what we already had to know in order to have the relevant concepts. For instance, what is the status of the following argument?

(P1) Moby Dick is a whale.
(C) Therefore, Moby Dick is a mammal.

Obviously, it doesn't hurt to explicitly add,

(P2) All whales are mammals.

It's true, and it makes the argument obviously valid. But was the argument valid without it? Is it part of the *meaning* of "whale" that whales are mammals? It is tempting to think so. In some sense you couldn't fully know what whales *really* are without knowing that they're mammals. On the other hand, P2 is a relatively recent discovery, before which, lots of people were capable of forming thoughts about whales. In *Moby Dick*, the narrator, Ishmael, believes that whales are fish. But surely he knows what the word "whale" means, even if he's badly mistaken about their nature. When someone shouts, "Avast! There blows a sperm whale off the starboard bow!" Ishmael understands perfectly well what's being claimed, and he forms the belief that there's a whale off the starboard bow. Someone, like Ishmael, could understand (P1) without knowing (P2). But *that means that (P2) does genuinely add*

some factual information to (P1). Thus, the argument is *invalid* without (P2) although valid with it. So, when in doubt, you should *always demand that the information be made explicit.* You might accidentally count some valid arguments as invalid, but in the end, no real harm comes from being too careful.

Box 1.3 Concepts and Definitions

To have the concept *square* or *cat*, etc. is just to have the ability to think thoughts about squares, cats, etc. Sometimes having a concept does seem to amount to knowing a definition: you have the concept *prime number* only if you know that it is something that is divisible only by itself and one. But having a concept isn't always a matter of knowing a definition: there isn't any particular thing you need to know about cats in order to think thoughts or understand sentences about cats. Knowing what they look like may be sufficient for having the concept; and knowing what they sound like may be sufficient; but neither is necessary.

In assessing arguments for validity, the background knowledge you're allowed to bring in is just the knowledge required for having the concepts involved in the premises and conclusions—that is, the background knowledge required for understanding these statements. Sometimes this may involve definitions, sometimes it won't.

Box 1.4 Definitions and Natural Language

The fact that dictionaries exist suggests that it's not hard to find definitions for all or most terms. But it's more complicated than this. Consider "bachelor." Suppose we define it as an unmarried male. That's not right. A 5-year-

old boy doesn't count as a bachelor. Fair enough, how about unmarried male of marriageable age? A man who was married, but has been widowed, satisfies this definition. Is he a bachelor? Some may say "yes"; some may say "no." Most will shrug their shoulders and look puzzled. What about a monk, or the Pope, who have taken vows not to marry? It doesn't seem that there's a clear-cut definition of even a simple concept like *bachelor*, but we can all think thoughts about bachelors. Surprisingly, it seems that having a concept is not, in general, the same thing as knowing a definition.

Notwithstanding potential disagreements over obscure cases, we all know that bachelors are, necessarily, unmarried men. So, even if the concept is not specified by a definition, it remains that the argument from John being a bachelor to John being unmarried is indeed valid. It's worth repeating, however, that when in doubt, it's best to err on the side of caution: if you make explicit the fact that all bachelors are unmarried as an added premise, then the argument will be obviously valid.

Exercises 1.3

A. **For each of the following arguments, say what background assumption would have to be added to render it a fully explicit, valid argument. Then say whether the argument was already valid without it.**

1 It's a domestic cat. Therefore, it's someone's pet.
2 It's a cat. Therefore, it's warm-blooded.
3 You're voting for Gil Fulbright. So, you're a Republican.

4 It's January. So, we'll get some snow soon.
5 It's Tuesday. Therefore, the day after yesterday is Tuesday.
6 You have Lyme disease. You must have been bitten by a tick!
7 All electrons are negatively charged. So, all electrons repel each other.
8 She was your date for your senior prom? I guess she really hated her parents!
9 I had an appendectomy six weeks ago. Therefore, I have undergone a surgical procedure in the past year.
10 You're down to a half a gallon of gas and you have 140 miles to go. Therefore, you won't get there without stopping to refuel.

While we're on the subject of meaning and its relation to validity, it's worth making note of a couple of important topics, to which we shall return in more detail later in the book.

Logical Terms

We've been looking at arguments whose validity hinges on the meanings of terms like "bachelor" and "circle," but arguments whose validity depends on the meanings of broadly ***logical terms***, like "some," "all," "and," "not," "if…then," are even more important. Consider the following:

(P1) All cetaceans are heterotrophs.
(P2) Mauyuk is a cetacean.
(C) Mauyuk is a heterotroph.

Even if you don't know what "cetacean" and "heterotroph" mean, or who Mauyuk is, you know what "all" means, and that's enough to know that this argument is valid.

We can show this a little more clearly by drawing a diagram. (P1) says that the class of cetaceans is included in the class of heterotrophs, that you can't be a cetacean without being a heterotroph. We can illustrate this by placing a circle that represents cetaceans inside a circle that represents heterotrophs. That way, anything that falls in the *C* circle is going to fall in the *H* circle as well:

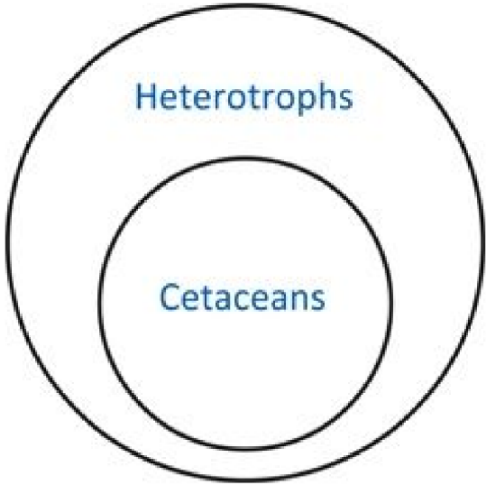


Figure 1.1

To represent the claim that Mauyuk is a cetacean, we use a dot or an x or something that stands for her, and we put it in the *Cetaceans* circle.

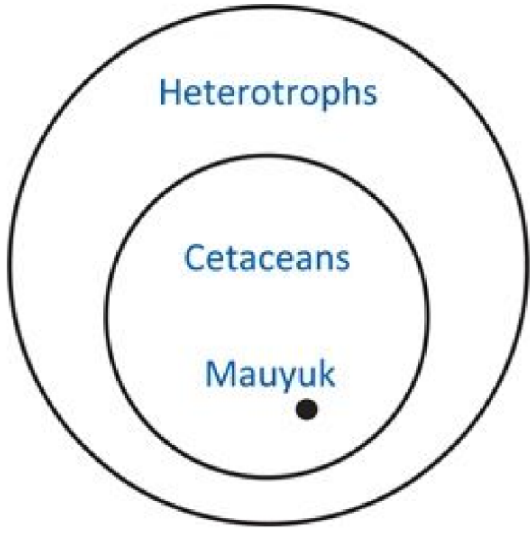


Figure 1.2

But now it is quite obvious that, because Mauyuk is a cetacean—she falls inside the smaller circle—she *must be* a heterotroph too—she must fall inside the larger, more inclusive circle. We can't draw a *Mauyuk* dot inside the *Cetaceans* circle without also putting it inside the *Heterotrophs* circle. This shows that from the fact that Mauyuk is a cetacean (P2) we can deduce with certainty that she is a heterotroph (C). On the other hand, if our second premise had instead claimed that Mauyuk was a heterotroph, we couldn't have validly inferred that she's a cetacean. This is because she might have—for all our premises have told us—fallen outside the *Cetaceans* circle but still inside the *Heterotrophs* circle.

As we progress, we'll see that we can learn a lot about validity by focusing on such logical terms: broad classes of arguments turn out to be valid precisely because of the patterns or forms in which such terms are used. The above case illustrates this. It just doesn't matter that the argument is about cetaceans or heterotrophs or about Mauyuk. All instances of the form

(P1) All Cs are Hs.
(P2) m is a C.

(C) m is an H.

will have to be valid. By contrast, arguments whose validity depends on specifics of the meanings of terms like "cetaceans" or "bachelor" are of little *general* interest.

Moreover, the meanings of logical terms are typically clearer and less ambiguous than the meanings associated with non-logical terms. Generally, when people say something of the form "*All* Xs are Ys" there is just one thing meant by the term "all." This also renders such terms as suitable targets in our attempt to characterize the properties of broad classes of arguments. By contrast, non-logical vocabulary is not generally so well-behaved. This brings us to the second meaning-related topic we need to briefly discuss: *equivocation*.

Equivocation

In judging the last argument valid, we reasonably assumed that the terms used in multiple locations ("cetacean," "heterotroph," "Mauyuk") meant the same thing each time they occurred. However, if a term is ambiguous, it can lead to the ***fallacy of equivocation***, where an argument uses a single word or phrase in two or more different ways, so that the argument has the appearance of being sound, even though it isn't. Here's a toy example to illustrate the concept. Consider the following argument:

(P1) My nephew is just a kid.
(P2) Kids are baby goats.
(C) Therefore, my nephew is a goat.

Obviously, "kid" is being used differently in these two premises. The argument looks valid, but it is only valid if "kid" means the same thing in both premises. However, if we use it to mean

"child," then (P1) is true but (P2) is false; if we use it to mean "baby goat," then (P2) is true but (P1) is false. So, on either reading, if the argument is valid, it's unsound; at least one of the premises is false. If, on the other hand, we use "kid" one way in (P1) and a different way in (P2), we can get two true premises, but now the argument's invalid, and again, unsound. We discuss equivocation in more detail in Chapter 10.

4. Logic and the Belief Bias

We've tried to be really careful in laying out what a valid argument is and how to evaluate simple arguments. However, even if you've taken everything we've said onboard, it is still easy to be misled about an argument's validity. Consider the following argument:

(P1) Anything that has a motor needs oil.
(P2) Cars need oil.
(C) Therefore, cars have motors.

Valid or invalid? Remember that validity is about conditional support and not about the actual truth values of the premises and conclusions. Many people are inclined to think the argument is valid. Now consider another argument:

(P1) Anything that has a motor needs oil.
(P2) Opprobines need oil.
(C) Therefore, opprobines have motors.

Most people are not inclined to think that this argument is valid.[1] Notice, however, that the two arguments are perfect parallels. We're inclined to think the first argument is valid because we know that the conclusion is true. The argument, however, is

invalid, as we can more easily see in the case of the second argument, where we immediately realize that things that don't have motors might also need oil for other reasons, and opprobines/cars might be among these things. This is a phenomenon known as the ***belief bias***: people tend to judge invalid arguments to be valid if the conclusion is something they already believe to be true.

Obviously this is a bad kind of mistake to make. Among other things, it will keep us stuck in a cognitive rut. If we once adopt a belief for no good reason at all, the mere fact that we now believe it will make bad reasons for that belief look like good ones. This tendency makes it difficult for us to learn, because it makes it difficult for us to be open-minded. And of course, it makes it difficult for us to be objective; we are naturally prone to find (real) flaws in the arguments of those who disagree with us, but to overlook (real) flaws in the arguments of those who agree with us. It's hard to overstate how bad this is from the perspective of believing all and only what's true.

The culprit here, of course, is System 1: we have an automatic, intuitive sense of validity and are prone to judge accordingly.[2] But this sense of validity is highly unreliable, because it's so heavily influenced by our fallible prejudgments about the conclusions. Even if we were perfect, infallible judges about the conclusions, this would still be a bad guide to validity, since it's quite possible for an invalid argument to have true conclusions or a valid argument to have false conclusions.

Our strategy for evaluating arguments is clear: Use System 2, don't let System 1 foist beliefs on you that just seem right. System 2 will be much more reliable, at least once you've learned how to properly distinguish between valid and invalid arguments.

Exercises 1.4

A. **Assess the following arguments for validity. Do**

so as quickly as you can, writing down whatever verdict pops into mind.

1. People who are opposed to freedom support gun control. Liberals support gun control. So, liberals are opposed to freedom.
2. People who don't care about the sick, the disadvantaged, and the elderly support cuts to welfare. Conservatives support cuts to welfare. So, conservatives don't care about the sick, the disadvantaged, and the elderly.
3. People who write graphic novels require the storytelling skills of a writer and the visual imagination of a good film maker. Film directors also need both of those skills. So, film directors write graphic novels.
4. People who totally buy into the scientific worldview are anti-religion. People who totally buy into the scientific worldview also believe in global warming. So, people who believe in global warming are anti-religion.
5. People who accept the results of well-established science accept the reality of global warming. People who accept the reality of global warming must be anti-religious. So, people who accept the results of well-established science must be anti-religious.
6. People with strong fundamentalist religious beliefs deny global warming. People who, for one reason or another, won't

honestly face the evidence deny global warming. So, people with strong fundamentalist religious beliefs won't honestly face the evidence about global warming.

7 Historically, the noble potato has long been a valuable food source in poorer societies across the world. Even today, in several countries it is common for potatoes to be on the dinner table almost every night of the week. As is well-known, this is true in Ireland, but also in Poland, Peru, and several other countries. So, we can conclude that these societies are still quite poor. They can still not be counted among the richer nations of the world.

8 Even people from the poorest of backgrounds have succeeded in life with hard work and dedication. So, anyone, even someone from a very poor background, can be successful.

B. **Go back through the arguments of section A just now, and this time, take your time evaluating them, keeping in mind the fact that we're all more likely to find an argument valid if we already believe the conclusion. Was there any difference between your quick assessment and your slower, more careful assessment?**

C. **Which of the following arguments are valid and which are invalid?**

> ***Watch out for belief bias, i.e., uncritically accepting an argument as valid, just because you think the conclusion is true.***

1 New York is bigger than Houston. Houston is bigger than San Francisco. So, New York is bigger than San Francisco.

2 Los Angeles is bigger than New York. Los Angeles is bigger than San Francisco. So, New York is bigger than San Francisco.

3 Abraham Lincoln and Bill Clinton were both U.S. presidents. Lincoln is dead, but Clinton is still alive. So, Lincoln was president before Clinton.

4 George H. W. Bush was the 41st president of the U.S. Bill Clinton defeated him in the presidential election in 1992. So, Bill Clinton was the 42nd president of the U.S.

5 The Empire State Building is in New York. You can see Brooklyn from the top of the Empire State Building. So, Brooklyn is in New York.

6 Mount Everest and K2 are both in the Himalayas. Mount Everest is taller than K2. K2 is the second tallest mountain in the Himalayas. So, Mount Everest is the tallest mountain in the Himalayas.

7 The Himalayas contain more tall mountains than any other mountain range on earth. Mount Everest is taller than any other mountain in the Himalayas. So, Mount Everest is the tallest mountain on earth.

8 Michael Phelps has won more Olympic gold medals than any other swimmer. Successful Olympic swimmers typically win more medals than other athletes. So, Phelps has won more Olympic gold medals than any other athlete.

9 Rio de Janeiro is in Brazil. Brazil is right beside Argentina. Argentina is in South America. So, Rio de Janeiro is in South America.

10 Washington D.C. is the capital of the U.S. The president's office, the Oval Office, is in the White house, which is in Washington D.C. So, the president lives in Washington D.C.

D. **Go back through the arguments in 1.1 B and re-assess for validity and soundness. Note any cases where belief bias initially leads you astray in assessing for validity.**

5. Why it Matters: Missing Premises and Insisting on Validity

Validity is the highest possible standard, since it demands that it's *logically* impossible that the conclusion is false and all the premises true. You might think that the very notion of logical possibility/impossibility is rather silly and impractical, especially when we remember that absurd situations, like talking typewriters, are logically possible. Do we really need such a high standard?

First of all, not all good arguments are valid, as we'll start to see in detail in Chapter 4. But valid arguments are the cleanest, simplest kind, and it's best to start where everything is pure and simple before moving on to the messy and complicated. Also, some arguments do meet this extremely high standard. If we didn't have the concept of validity, we wouldn't be able to explain how these

arguments differed from other arguments, with a weaker degree of conditional support.

Most importantly, when we insist on validity, this forces us to make explicit premises that we had left unstated. This, in turn, forces us to directly confront our unarticulated and maybe unconscious assumptions. This can have an enormous beneficial effect on our thinking.

Suppose we had a less exacting standard, one that accepted the original Ice argument as perfectly adequate. We would have a large blind spot in our understanding of the world. Someone who thinks the reasoning of that argument is adequate has just *ruled out*, without giving it a thought, the possibility that there might be different types of water ice with different melting points. If we do not require that it is *logically* impossible for the premises to be true and the conclusion to be false, we will *fail to track the truth* in an important way: we will implicitly assume we know things that we don't—in this case that *all* solid water melts at 32° F. That is why we hold arguments to the highest possible standard. If we don't, we will have intellectual blind spots.

Insisting on validity, rejecting a line of reasoning until it's clearly valid, compels us to insert the extra premise:

(P2) Solid water will always melt when heated above 32° F.

With this assumption out in the open, we no longer have that intellectual blind spot. We can now assess this premise and recognize that our evidence for it may indeed be surprisingly weak —yes, any water *I have seen* melts when heated above 32° F, but is it *obvious* that even in unusual or exotic circumstances that is *always* the case? Surely not. Answering that question demands scientific research, not just casual observation of the behavior of water under normal circumstances.

An incomplete argument is called an ***enthymeme***. Converting enthymemes into fully stated arguments is a way to drag unstated assumptions into the light. Very often, the unstated assumptions,

the missing premises, are the weakest. In many cases, they turn out to be obviously false, or at best highly controversial and unsupported by the evidence. The belief bias means we often won't notice this unless we actively engage System 2 and self-consciously assess the stated argument's validity. So, insisting on validity, and the associated reconstruction of arguments and interrogation of the premises we uncover is one of the most powerful tools for reasoning and arguing reliably. We'll have more to say about enthymemes and about argument reconstruction in Chapter 3, but for now, we want to note that enthymemes are very common, and that filling in the missing premises to make an argument valid is often extremely illuminating.

Through much of history, it was accepted that the earth sat entirely stationary at the center of the universe, with the sun and other objects orbiting around it. When scientists and philosophers considered the possibility that the earth orbited the sun and that it also rotated on its own axis, they were met with great skepticism. Just the rotation of the earth on its axis seemed to be ruled out by an argument from the most straightforward observational data:

> (P1) If the earth rotates, we are moving at about 1,000 mph.
> (P2) We don't *seem* to be moving.
> (C) So, the earth isn't rotating.

This seems pretty reasonable, especially if you're living in the 1500s and have the belief bias working in favor of the argument, rather than against it. It was known that the earth's circumference was about 24,000 miles, and so, given that the earth has to rotate once every 24 hours, someone at the Equator would be moving at 1,000 mph. Even far north of that, you could work out that we would be moving at a substantial fraction of that speed. So, (P1) is true. However, the argument is not valid, and we can easily see what's missing. To be valid it needs a further premise:

> (P3) If we are moving at about 1,000 miles per hour, it will