5

Conditional Reasoning Part I

Three Kinds of Conditionals and the Psychology of the Material Conditional

5.1 Introduction

We know that A-imagining involves contemplating possibilities in a rich and epistemically safe way. But there are many things it can be to "contemplate possibilities." In this chapter and the next, I want to focus on one particularly important kind of (often) rich or elaborated thought process during which we contemplate possibilities: *conditional reasoning*. Episodes of conditional reasoning, as I will understand them, are thought processes that result in judgments with an if-then structure. We can reason conditionally about what is or will be the case, given that certain other things are or will be the case—judging, for instance, that *if Henry is at the meeting, then he looking at his phone*, or that *if I go to the block party, then I will try the bean dip*. And we can reason conditionally about what would have been the case, had things been different—concluding, for example, that *if I hadn't had the bourbon, I'd feel better today* or that *if Clinton hadn't used a private email server, she would have won the election*. In linguistically expressing such thoughts, we use if-then statements, otherwise known as conditionals. One of the central hopes for a theory of imagination is that it sheds some light on our capacity to reason with and about conditionals.

I want to set reasonable goals, however. Conditionals generate deep and continuing controversies within both philosophy and psychology. There is currently little consensus in how their truth conditions are to be analyzed, nor in whether all types of conditionals have truth conditions at all. (See Bennett (2003) and Edgington (1995) for lucid overviews of these debates.) Nor is there consensus in experimental psychology concerning the psychological states at work in their appraisal (see Evans & Over (2004) and Johnson-Laird & Byrne (2002) for discussion), or even in the normative standards that ought to apply to judgments about conditionals in experimental contexts (Oaksford & Chater, 2003). My limited goals are twofold: first, to articulate a coherent view—a *how plausibly* story—on which inferring and reasoning with conditionals draws only upon beliefs. (Some of these beliefs will involve mental images as proper parts (see Chapter 4).) My second goal is to show that, correct or not, such a view is preferable to any that posits *sui generis* imaginative states. Doing this much will serve the book's broader

Explaining Imagination. Peter Langland-Hassan, Oxford University Press (2020). © Peter Langland-Hassan.
DOI: 10.1093/oso/9780198815068.001.0001

purpose of explaining imagination, insofar as it shows how the A-imagining that occurs during conditional reasoning can be reduced to patterns of inference involving more basic folk psychological states (beliefs, primarily); this reduction undermines whatever attraction views positing *sui generis* imaginings may seem to have. The questions I leave open—such as the proper analysis of the semantic difference between indicative and subjunctive conditionals, or the mental states that are *in fact* exploited during conditional reasoning—are the kinds of questions that can only be answered by a formal philosophical or psychological theory of conditionals, both of which fall beyond the scope of this book. It will, however, be important to understand the key *questions* in this area, including those surrounding the distinction between indicative and subjunctive conditionals, and the relation of each to the material conditional of formal logic. To date, these crucial distinctions among kinds of conditionals have, for the most part, gone ignored in philosophical discussions of imagination's relation to conditional reasoning.[1] As we'll see, grasping them is essential to understanding the role of imagination in conditional reasoning, and in "modal epistemology" more generally.

5.2 Modal Epistemology?

Before diving in, a few words on the project of *modal epistemology* in general. The questions of modal epistemology are questions of how we arrive at knowledge of possibilities and necessities. These questions differ, however, depending on the sense of 'possible' and 'necessary' in play. At times, we use the term 'possible' in an epistemic sense, to mark what is not ruled out by what we know. For instance, it is possible, for all I know, that 237 × 345 is 84,425. Nothing I believe rules it out. However, I have now carried out the calculation and see that the answer is 81,765. It is now no longer epistemically possible, for me, that 237 × 345 is 84,425. Knowing what is and is not possible in this sense has only to do with knowing our own beliefs; there is nothing especially puzzling about our knowledge of these kinds of ("merely epistemic") possibilities.

Other uses of modal terms have a more objective air. There is another sense in which, even before I did the calculation, it was *not* possible that 237 × 345 is 84,425. In the realm of mathematics, this more objective form of possibility is referred to as *logical* or *conceptual* possibility. But I will follow more recent convention in using the term *metaphysical possibility* to mark the entire realm of objective possibilities. (I will leave open the question of whether there are logical

[1] The most detailed existing discussion I know of is in Williamson (2007, pp. 134–55), though Williamson focuses almost entirely on subjunctive counterfactual conditionals to the exclusion of indicatives and the material conditional.

or conceptual possibilities that are not also metaphysical possibilities, as those controversies won't touch the questions at issue here.) Unlike the multiplication example, some propositions that are not epistemically possible (for me) nevertheless remain *metaphysically possible* in the sense that they *could have* been true. Suppose, for instance, that I count the objects on my desk and see that they are five. It is now *not* epistemically possible for me that there are ten objects on my desk; there being ten is not compatible with what I know. However, it remains metaphysically possible that there *could have been* ten objects on my desk. That is a way the world could have been, but is not.

It is a good question how we know that certain things *could have* been the case—in this more objective, metaphysical sense—given that we know they did not occur. How do we figure it out? Since perception doesn't seem to the point, imagination is often pushed onstage to answer. It is said, following Hume (1738/2012), that imagination is to the possible as perception is to the actual. Perhaps it is not obvious *how* imagination would offer us a window onto the possible-but-not-actual. Unactualized possibilities do not, after all, causally impinge upon our imagination in the way that ordinary perceived objects impinge upon our senses. But, nevertheless, some find it highly intuitive that imagination plays such a role; and they may be satisfied to show that there is nothing incoherent in the idea that imagination (or an idealized version thereof) offers us reliable access to facts about the possible (Yablo, 1993; Chalmers, 2002; Kung, 2010).

An alternative way to approach the question of how we determine what is possible is through examining conditional reasoning. When we say that Hillary Clinton could have won the 2016 election, we typically have in mind certain counterfactual conditions under which she would have won. For instance, if Clinton had not used a private email server, we may reason, she would have won the election. Likewise, if Earth had been struck by a wave of giant asteroids millions of years ago, we may think, it would now be devoid of life. Answering how we come to know the relevant facts about what *would* have happened—and thus what is metaphysically possible but not actual—looks to be part and parcel with coming to know related conditionals. Typically, when we decide that if *p* then *q*, we have found *q* to be metaphysically possible, in the sense that *q* could have happened. (Note the formal difference, however: it is one thing to say "if *p* had been the case, *q* would have been," quite another to say, "it could have been that *q*.") Some have indeed argued for a tight connection between metaphysical modality and counterfactual reasoning, proposing that our knowledge of the former relies entirely on our ability with the latter (Williamson, 2005, 2007, 2016). While I am sympathetic to that project, I won't defend it or rely on it here. (See Strohminger & Yli-Vakkuri (2017) for more on this debate.) I will instead limit my discussion of modal epistemology to querying the role of imagination in our reasoning with and about conditionals. This includes not only counterfactual subjunctive conditionals ("if *p* had been the case, then *q* would have been"), but forward-looking hypotheticals in the indicative mood as well, such as: *if John mocks the dean, he will be sorry*, or *if Julia goes on a cruise, she will regret it*. These, too, are judgments about merely possible situations.

Focusing on conditional reasoning may leave my treatment of modal epistemology incomplete by the lights of those who reject an equivalence between conditional reasoning and judgments about possibility and necessity. I think this merely amounts to a difference in aims and interests, however. My aim is to explain the platitudinous facts about imagining that anyone needs to accept—e.g., that it guides pretense, helps us to plan our actions, is used when we make ordinary judgments about what could have (but didn't) happen, underlies creativity and our engagement with fiction, and so on. Explaining imagination's role in conditional reasoning serves that end. Some others who theorize about imagination when doing modal epistemology—such as Chalmers (2002), Yablo (1993), and Kung (2010)—seem to have a different project. They take, as common ground, a set of modal claims—many inspired by the work of Putnam (1975) and Kripke (1980)[2]—and aim to describe a kind of mental process (an idealized form of *imagination*) that could be relied upon (or not) to ground our knowledge of those claims. As these modal claims are *prima facie* surprising to most, explaining our knowledge of them requires considerable revision (and idealization) of the commonsense notion of A-imagining.

I do not pass judgment on that project; I just set it to the side, for the purposes of this book. The question of how and why we infer the conditionals relied upon in everyday life—and the role of imagination in our doing so—is difficult enough. It is a question we need to answer whether or not there are such things as *a posteriori* necessities, and irrespective of any controversial claims about *particular* possibilities (such as that we could have physical duplicates who lack phenomenal consciousness (Chalmers, 1996)). Explaining the imaginings that occur during conditional reasoning is essential to explaining how we get about in the world. That is modal epistemology enough. Maintaining this focus has the added benefit of tethering our inquiry to related literatures on conditionals in the philosophy of language and experimental psychology.

5.3 Conditionals: Metaphysics and Psychology

To begin, we need to distinguish two different, if related, questions we might ask about conditionals. First, we can ask the metaphysical question of what conditionals *are*. Here we are asking for a theory of conditionals themselves. Typically, a theory of conditionals tries to explain what conditionals are by giving a systematic

[2] These include surprising "*a posteriori* necessities," such as that water is necessarily H20, and that Hesperus is necessarily Phosphorus.

account of their truth conditions or semantics. This involves comparing and contrasting different kinds of conditionals with respect to their ability to fit a certain framework for understanding truth conditions (or, alternatively, appropriateness conditions for their utterance). The second question we can ask concerns the nature of the psychological states exploited when reasoning about conditionals, or when coming to infer one conditional as opposed to another. This psychological question is my focus here, as imagination is most naturally, and most commonly, invoked in explanations of how we reason our way to conditional beliefs, as opposed to in theories of their semantics and truth conditions.

Traditionally, most philosophical discussions of conditionals have focused on their metaphysics (Bennett, 2003; Edgington, 1995; Lewis, 1973; Lycan, 2001; Stalnaker, 1968), with experimental psychologists attending more to the nature of the mental processes exploited in assessing and inferring conditionals (Byrne, 2007; Evans & Over, 2004; Johnson-Laird & Byrne, 2002). Yet there is always interplay between the two questions. Philosophical theorizing about the truth conditions of conditionals treads deeply into psychological questions concerning how and why we accept the conditionals that we do. And psychological theorizing about the nature of the mental states exploited in conditional reasoning, and related experimental designs, are inevitably tied up in assumptions concerning the proper semantics for conditionals—assumptions according to which some participant responses to prompts are *mistakes* in need of explanation (Johnson-Laird & Byrne, 2002; Oaksford & Chater, 2003).

In the next section, I canvas some reasons commonly given for distinguishing three different types of conditionals—material, indicative, and subjunctive conditionals—by their different associated truth conditions. If these different kinds of conditionals indeed show systematic differences in their truth conditions—if there are different things that *if-then* means in each case—we can expect those differences to show up in whatever imaginings are at work in generating our beliefs in conditionals. One of the main projects of this and the following chapter is to show how imagination needs to be understood somewhat differently when theorizing about each type of conditional. In each case, I will argue, a reductive account is available.

5.4 The Material Conditional and Its Relation to Indicative and Subjunctive Conditionals

In all the foggy terrain surrounding conditionals, there is nothing clearer than the metaphysics of the material conditional, familiar to systems of formal logic. The material conditional is defined in terms of a simple truth table. Letting the horseshoe ("⊃") stand for the relation of material implication, the truth of '$p \supset$ 'q is a simple function of the truth of p and q. That gives us four possibilities: both p and *q* are true; *p* is true while *q* is false; *p* is false, while *q* is true; and both *p* and *q* are false. The truth table used to define the meaning of '$\supset$' tells us that '$p \supset q$' is true in all these situations save where *p* is true and *q* is false. A material conditional's truth, then, is entirely a function of the truth or falsity of the two propositions flanking the horseshoe. This is the sense in which the material conditional is *truth functional*.

Many of the philosophical puzzles concerning conditionals spring from the observation that what goes for the material conditional does not obviously apply to the conditionals of natural, spoken languages. I will follow convention in distinguishing two classes of conditionals that occur within natural language: indicatives and subjunctives. Indicative conditionals are marked by the indicative mood ("is," "will") in the manner of: *If John is at home, then he is studying*. Subjunctives are marked by the subjunctive mood ("had" or "would") and comprise conditionals such as: *If John had studied, then he would have passed the exam*. The study of "counterfactuals" in philosophy focuses on subjunctive conditionals in the past tense, with formulations such as "If he had dropped the rock, it would have broken his toe." The person asserting such a counterfactual typically doesn't believe the antecedent conditions to hold. However, the subjective mood and the notion of an antecedent thought to be counterfactual don't always march in lockstep. There is, for instance, the doctor who in diagnosing malaria comments: "If he had contracted malaria, these are exactly the symptoms we would expect." For reasons we will come to, it is nevertheless customary to theorize about subjunctive counterfactuals as a class, distinguishing them from indicative conditionals in both the present and past tenses.

Most contemporary theories of conditionals deny that the indicative and subjunctive conditionals of natural language are truth functional in the manner of the material conditional (Bennett, 2003; Edgington, 1995; Lewis, 1973; Lycan, 2001; Stalnaker, 1968). The reasons are easy to see when we observe that the truth table for the material conditional '$p \supset q$' is identical to that for the disjunction: not-*p* or *q* (i.e., $\sim p \vee q$). That is, '$p \supset q$' is true in every situation where '$\sim p \vee q$' is true, and false in every situation where '$\sim p \vee q$' is false. This makes plain that the mere falsity of *p*, or the truth of *q*, is each sufficient for the truth of '$p \supset q$.' Yet, in the case of subjunctive counterfactual conditionals, the antecedent is almost always false, and thought to be false by the person uttering it. If the truth conditions for subjunctive conditionals are the same as those for the material conditional, then almost all counterfactual conditionals must be true. This clashes badly with our actual use and appraisal of subjunctive counterfactual conditionals. We are not, for instance, apt to judge *both* of the following counterfactuals true simply because they have false antecedents:

(A) If Clinton had not used a private email server, she would have defeated Trump

(B) If Clinton had not used a private email server, Trump would still have won.

Pundits clash over which of (A) or (B) is true. But none argue that both are true. So subjunctive counterfactual conditionals appear not to be truth functional in the manner of the material conditional. Some other analysis of their truth conditions is needed.

Similar problems plague attempts to equate indicative conditionals with the material conditional. Let '$\rightarrow$' stand for the indicative conditional relation, such that '$p \rightarrow q$' is an arbitrary conditional in the indicative mood. If '$p \rightarrow q$' has the same truth conditions as '$p \supset q$,' then '$p \rightarrow q$' is true whenever p is false and whenever q is true. But, intuitively, in assessing whether '$p \rightarrow q$' is true, we wish to know more than whether p is false or q is true. We want to know whether a certain connection holds between p and q.[3] This tension comes to the fore in the "paradoxes" of material implication. To pull an example from Stalnaker, if the indicative conditional is logically equivalent to the material conditional, then the following should be a valid argument: "the butler did it; therefore, if he didn't, the gardener did" (1975, p. 136). For the premise ("the bulter did it") is in effect the negation of the antecedent of "if he didn't, the gardener did"; and we know, from the truth table for material implication, that $p \rightarrow q$ is true whenever p is false. Similarly, according to the logic governing the material conditional, any false conditional must have a true antecedent. Staying with Stalnaker's example, it seems absurd to propose that, in denying the conditional "If the butler didn't do it, the gardener did," we must thereby accuse the butler.

Efforts have nevertheless been made to defend the idea that indicative conditionals share the material conditional's truth functionality. It can be replied, for instance, that the oddities we see in the above "paradoxes" are pragmatic in nature and don't touch the logical validity of the inferences (Grice, 1989; Jackson, 1987). In situations where one already knows that not-p, for instance, it is conversationally *inappropriate* to say "$p \rightarrow q$." For there is an implication carried by utterances of $p \rightarrow q$ that the speaker is in doubt as to whether p, just as there is an implication carried by utterances of "p or q" that the speaker is in doubt as to both p and q. This flouting of a pragmatic norm might be thought to explain away the sense that it is logically invalid to infer that $p \rightarrow q$ from not-p. The reason it seems "off" to infer $p \rightarrow q$ from not-p is that it is inappropriate to *say* $p \rightarrow q$ when one already believes that not-p—inappropriate because it would have been more informative for the speaker to have instead said that not-p, just as it is more informative to say that p, instead of p or q, when one believes that p.

Few have been persuaded by such efforts, however. As Dorthy Edgington observes, the problems with analyzing indicative conditionals on a par with the material conditional occur at the level of rational inference, irrespective of any norms to be succinct in conversation. This is particularly evident in the many

[3] Strawson (1986), for instance, analyzes the meaning of $p \rightarrow q$ as, roughly, "There is a connection between p and q that ensures that: $p \supset q$."

everyday contexts where one is confident in the truth of some proposition but lacking in absolute certainty. To take Edgington's example, she is confident that her husband is not home; and she is also confident in the conditional: if he is at home, he will be worried about my whereabouts (because she is working later than usual). Edgington is also confident that the Queen is not at home. Yet there is nothing irrational in her *rejecting* the claim that if the Queen is home, then she will be worried about my whereabouts. On the truth-functional analysis of indicative conditionals, however, confidence in not-*p* ("The Queen is not at home") should always warrant equal confidence in "$p \rightarrow q$." "We need to be able to discriminate believable from unbelievable conditionals whose antecedent we think false," Edgington explains. "The truth-functional account does not allow us to do this" (1995, p. 245).

Another point against equating '$p \rightarrow q$' with '$p \supset q$' traces to differences in whether contraposition succeeds for each. For the material conditional, '$p \supset q$' and '$\sim q \supset \sim p$' are equivalent—that is, true in all the same situations. Yet this isn't always the case with indicatives. Here is an example from Jonathan Bennett: "I accept that even if the Bible is divinely inspired, it is not literally true; but I do not accept that if it is literally true, it is not divinely inspired" (Bennett, 2003, p. 30). Other procedures that are valid for the material conditional but of questionable validity for indicatives include Transitivity and Antecedent Strengthening.[4] But, more generally, the idea that indicative conditionals can be equated with the material conditional, with counterexamples explained away as merely pragmatic, has suffered from the availability of a quite different and *prima facie* more attractive alternative for understanding their nature: the Ramsey test.[5] We will look closely at the Ramsey test in Chapter 6.

I have so far canvassed just a few reasons for distinguishing subjunctive and indicative conditionals from the material conditional. The consensus on these matters is that we cannot tether our investigation of conditional reasoning entirely to principles appropriate to the material conditional. For not all procedures that are valid with respect to the material conditional extend to indicative and subjunctive conditionals. If indicative and subjunctive conditionals are the conditionals of everyday life, then our main interest here should be in the psychological states that allow us to evaluate them. Nevertheless, in the balance of this chapter, I want to focus specifically on the material conditional, asking: if and when we reason in accordance with the logic governing the material conditional, what sort of mental states must we exploit? For even if we do not always, or even typically,

[4] The truth table for the material conditional guarantees that if $p \supset q$ is true, then so will be $p \,\&\, r \supset q$. This does not always hold for indicatives. True: If you jump out of an airplane from 3,000 feet in the air, you will perish. False: If you jump out of an airplane from 3,000 feet in the air and pull the ripcord on your parachute, you will perish.

[5] Though see Jackson (1987) for an attempt to fold the key insights of the Ramsey test into an account that still equates the truth conditions of '$p \rightarrow q$' with '$p \supset q$.'

treat the indicative and subjunctive conditionals of natural language as we would if they had the truth conditions of the material conditional, we might in some cases. Moreover, given the historical interest of philosophers in systems of natural deduction—wherein the material conditional resides—some will no doubt hold out hope for understanding either indicative or subjunctive conditionals by close, if not perfect analogy to the material conditional. So it will be worthwhile to consider the relation between the material conditional and our own psychological states before moving on, next chapter, to focus squarely on the nature of indicative and subjunctive conditionals. In particular, we need to ask whether we have good reason to posit *sui generis* imaginative states as a means to explaining our relationship to the material conditional.

5.5 The Material Conditional and Assumptions: Conditional Proof and *Reductio ad Absurdum*

Suppose that we are presented with the conditional statement "if *p* then *q*" and asked to assess its truth. Intuitively, when we consider whether "if *p* then *q*" is true, we imagine that *p* and see if *q* is also true, or at least likely, given that *p*. This is, at least, one thing we might do. But suppose, further, that in answering we must limit ourselves to deductively valid inferences in keeping with the logic of the material conditional and formal principles of natural deduction more generally. (Of course, we are not so limited in everyday life, where inductive and abductive inferences are also available; the point for now is to focus on how imagination relates to procedures within formal systems of natural deduction.) What can the imagining that occurs in considering "if *p* then q" amount to, if we are limited to psychological states and processes that mirror the steps and inferential principles of natural deduction?

We should first observe that not all assessments of "if *p* then q" will require anything that intuitively seems like imagining that p. If, for instance, we already happen to believe that not-*p*, or that q—or to believe other things from which we can deduce that not-*p*, or that q—we can immediately infer that "if *p* then q" is true, relying solely on inferences among our beliefs. This follows straightforwardly from the truth table by which the material conditional is defined. Likewise, if we happen already to believe that *not-p or q*, we will be warranted in believing *if p then q*. On the other hand, if we already believe that *p and not-q*, we will have immediate reason to reject the conditional. In all these cases, deciding whether to believe the material conditional did not require us to step outside of our beliefs to arrive at a judgment about what would happen if *p*. Intuitively, it did not require us to imagine that *p.*

But what about cases where we have no idea whether *p* or *q* and where we lack any beliefs on the basis of which we might deduce either's truth or falsity? How then might we decide whether to believe "if *p* then *q*," limiting ourselves to psychological states that mirror the steps of a deductively valid proof in formal logic? Could imagination get a foothold here? Are there times when we imagine that *p* so as to see what would follow (deductively) from *p*? Here the method of *conditional proof* suggests itself. Within a system of natural deduction, a conditional proof can be used to assess whether the conditional "A ⊃ C" follows from a set of premises, when A and C are not themselves premises of the proof. A simple example occurs in the context of proving that the hypothetical syllogism (also known as transitivity of implication) is a valid form of inference—i.e., that it is always truth-preserving to infer "A ⊃ C" from "A ⊃ B" and "B ⊃ C." Using the method of conditional proof, we can "assume" the antecedent of the conditional in question and see if, in conjunction with our other premises, we are able to derive the consequent of the conditional. If we are, this serves to prove that the conditional itself follows from the (not-assumed) premises. The assumption of A, together with any steps following it in the sub-proof it initiates, are "discharged" at the conclusion of the proof, insofar as they are not put to use in any further derivations.

On paper, the conditional proof of the hypothetical syllogism looks like this:

1: A ⊃ B (premise)
2: B ⊃ C (premise)
3: A (assumed for the purpose of conditional proof)
4: B (from 1 and 3)
5: C (from 2 and 4)
6: A ⊃ C (from lines 3, 4 and 5, by conditional proof; steps 3–5 are discharged)

Here we have a deductive method for arriving at a new conditional, when the antecedent and conclusion are not among the non-discharged, non-assumed premises. If we were to *psychologize* this procedure, understanding each step as a token mental state in a sequential process of reasoning, we could ask what *kind* of state corresponds to each step. A natural picture suggests itself: Steps 1 and 2 correspond to beliefs—beliefs on the basis of which we wish to know whether we may infer the conclusion in step 6. Steps 3, 4, and 5, on the other hand, might be thought to correspond to mere suppositions or *sui generis* propositional imaginings (or, indeed, "assumptions"). For they serve to register the (mere) assumption of A and the determination of what follows from that assumption, given the two not-assumed premises. The fact that the assumed steps are later discharged—and thereby, in a sense, *quarantined* from any further derivations—meshes with the idea that their contents are not preserved among one's beliefs; it also seems to fit with the psychological platitude that we can imagine, suppose, or assume things we do not believe.

Another context where psychologizing the procedures of a system of natural deduction seems to suggests a role for imaginative states are proofs via *reductio ad absurdum*, where we "assume true" a particular proposition in order to show that its acceptance would lead to a contradiction. When constructing a *reductio* of *p*, we typically will not believe that *p*, after all. Thus, were we to carry out a *reductio* "in our heads"—assuming that *p* in the process—it may seem that we must exploit a mental state *other than* one of our beliefs to record the assumption. Here again we seem to have reason to posit *sui generis* imaginative states, corresponding to the "assumed" step(s) in a *reductio*.

5.6 Psychology and Systems of Natural Deduction

We are considering whether the methods of conditional proof and *reductio ad absurdum* within systems of natural deduction give reason to posit *sui generis* imaginative states. A first question to ask about these cases is whether, in order to arrive at the kinds of conclusions allowed by conditional proof and *reductio ad absurdum*, we *must* exploit mental states that correspond in a roughly one-to-one way to the steps of such proofs—positing something like *sui generis* "imaginings" or "supposings" wherever an assumption appears. Is it in some sense a priori, or *necessary*, that such mental states are used to arrive at the all the deductively valid inferences that we in fact can make? Here the answer is a clear no. Prior to the development of systems of natural deduction in the 1930s (Gentzen, 1934; Jaskowski, 1934), "axiomatic" systems of formal deduction—such as those developed by Frege (1879) and Ackermann & Hilbert (1928)—held sway. All the theorems that can be proven in systems of natural deduction can be proven in axiomatic systems as well.[6] Yet axiomatic systems don't employ "assumptions" within proofs at all; instead they contain a set of basic assumptions (the axioms) and, typically, employ a single inference rule of *modus ponens*. If we were to psychologize each step of an axiomatic proof—mapping it to a particular mental state—each step could be seen as corresponding to a suitable belief, insofar as each step is either a premise or an axiom. Sound axiomatic systems can do all the same work as sound systems of natural deduction that employ assumptions; indeed, many logic textbooks teach a method for converting proofs of one kind into proofs of the other.

However, axiomatic proof systems are notoriously complex and unwieldy. A fairly straightforward proof within a system of natural deduction may require dozens of steps in an axiomatic system—with each step containing long strings of linked propositions. The byzantine complexity of axiomatic systems is one reason more "natural" systems were sought. One might argue, therefore, that axiomatic proof systems do not represent a *plausible* alternative to systems of natural

[6] Pelletier (2000) provides a helpful history of the transition from axiomatic systems to systems of natural deduction, and the spread of the latter throughout North America in the 1950s and 1960s, thanks to its adoption in influential logic textbooks.

deduction, provided that we are looking for models that may closely correspond to our actual psychological processes when evaluating conditionals.

While the objection has merit, it would be an error to conclude from the difficulties we may have in understanding a theory that specifies the nature of our cognitive workings that we do not, in fact, make use of the states or processes that the theory describes. The challenges one may experience in understanding the nature of artificial neural networks, or of Bayesian updating, do not, for instance, tell against the hypothesis that our minds exploit processes well modelled by artificial neural networks, or that calculate Bayesian probabilities. We may reason in accordance with principles, or through the use of computational processes, that we are in no position to articulate and that we would struggle to comprehend. So much is a working assumption in computational cognitive science. The comparative complexity of axiomatic systems does not, then, render them otiose as hypotheses about our cognitive underpinnings.

Now, I do not, as it happens, think that any axiomatic proof system *is* a faithful model of our thought processes. They were never created to be such. Nevertheless, they serve as a helpful reminder that there is in principle no difficulty in the idea that deductive inferences—including inferences in favor of new conditionals—can take place without the use of mental states whose contents mirror those of the assumptions or suppositions in systems of natural deduction.

Such reminders aside, we can still ask the more pressing question of whether the theory that we exploit mental states mirroring the steps and inferential rules of systems of natural deduction remains a *plausible* theory. Is there good reason to think that, when evaluating conditionals, we enter into mental states that mirror the steps of proofs within natural deduction—assumptions (or "suppositions") included? Here again the answer is no. As Jonathan St. B. T. Evans comments in a review article summarizing the last forty years of psychological research on human reasoning, "few reasoning researchers still believe that [deductive] logic is an appropriate normative system for most human reasoning, let alone a model for describing the process of human reasoning" (2002, p. 978). While in the late 1960s, it was still common among psychologists (under the influence of Piaget) to hold that adult human thought processes unfolded in ways that mirror the steps of deductive logical proofs, "it soon became apparent that…participants performed very poorly" on abstract deductive reasoning tasks aimed to reveal those capacities (p. 980). These tasks were specifically devised to abstract away from potentially biasing contextual information, so as to allow participants to focus on logical structure. In a typical example, participants are asked to assess whether "not-A" must follow from "If A then B," and "not-B." It is now common coin among psychologists studying human reasoning that people make widespread and systematic errors in their judgments concerning the validity of different forms of deductive inference (see Manktelow (1999) for a review).

Perhaps the most famous and robust paradigm in this literature—the Wason selection task (Wason, 1968)—concerns the evaluation of conditionals. In a typical version of that task, participants are shown four cards and told that each has a letter on one side and a number on the other. Only the letter side is visible on two cards, while the only number side is visible on the others. Participants are asked which cards they would need to overturn in order to evaluate the truth of the conditional: "If there is an even number on one side of a card, there is a vowel on the other." The interesting—and very robust—result is that over 90 percent of participants fail to suggest turning over a card that shows a consonant, despite the fact that, should there be an even number on the other side of that card, the conditional is falsified. One way to put the apparent implication is that, when evaluating conditionals, most people fail to consider the relevance of situations where the consequent is falsified. And yet, if people were hard-wired to reason in accordance with the principles of natural deduction, it is hard to see why they should so often fail to recognize the importance of such situations to the truth of the conditional they are to evaluate.

In another well-known and equally robust result from this literature, while participants reliably affirm that *modus ponens* (if *p* then *q*; *p*; therefore, *q*) is a valid form of inference, only about 60 percent of undergraduate university students answer that *modus tollens* (if *p* then *q*; not-*q*; therefore, not-*p*) is valid (Evans, Newstead, & Byrne, 1993). Further, participants frequently endorse fallacies, such as "denying the antecedent" (viz., "if *p* then *q*; not-*p*; therefore, not *q*") and, especially commonly, affirming the consequent (viz., "if *p* then *q*; *q*; therefore, *p*") (Evans, Clibbens, & Rood, 1995). Such results have spurred psychologists to posit psychological processes that would explain them. These processes have properties at odds with systems of natural deduction, insofar as they are specifically designed to explain the ways in which human judgments systematically diverge from the patterns allowed by systems of natural deduction. One of the most influential proposals of this sort is Johnson-Laird's (1983) and Johnson-Laird & Byrne's (2002) "mental models" hypothesis, which we will consider in some detail below.

For now, two important points can be made in summation. First, there is no necessary entailment that our thought processes, when evaluating conditionals, mirror the steps of a system of natural deduction—including its use of "assumptions." When we follow along with a request to "assume that *p*"—be it in ordinary conversation, or when assembling a *reductio*—there are a variety of things we may be doing that are not entering into a *sui generis* mental state of assuming, supposing, or imagining with *p* as its content. We saw that there are well-developed alternative logics that incorporate no assumptions. While there is no reason to think that these systems describe our thought processes as they actually occur, they serve as exemplars for the in-principle dispensability of "assumptions" and "suppositions" at the level of cognitive processing. (We will consider other possibilities of this sort below and in the chapter to follow.) Second, systems of natural deduction are not, in general, descriptively adequate with respect to ordinary human reasoning. While it certainly *could be* that we nevertheless, at times, make use of *sui generis* imaginative states when evaluating conditionals, the important role that assumptions play in systems of natural deduction give us no reason to posit such states, for the simple reason that such proof systems are not themselves descriptively adequate with respect to human reasoning.

5.7 Conditional Proof and *Reductio ad Absurdum* Revisited

And yet, even if systems of natural deduction are not descriptively adequate with respect to human reasoning, one might think that some *pieces* of them are, some of the time, for some people. In particular, if one can't really see how to do without assumptions when deductively inferring a conditional, or conducting a *reductio*, it may be tempting to hold on to the idea that assumptions play a role in the mind comparable to the role they play in natural deduction. For this reason, it will be worthwhile to show how the methods within natural deduction that seem to cry out for mental states of "assuming" (or "supposing" or "imagining") can be reconceived so as to involve only belief. We've already seen that such reframings are possible, in principle, by reflection on axiomatic proof systems. However, knowing that assumptions are eliminable in principle may leave one skeptical that they can be avoided in practice. Thus, in the examples below, I will limit myself to mental states and inference rules that, like those of natural deduction, translate smoothly to the terms of ordinary folk psychology. This will help to reinforce the point that any apparent practical need for cognitive equivalents to "assumptions" (via *sui generis* imaginative states) is illusory.

5.7.1 Conditional Proof without Assumptions

Let's return, first, to the assumptions within a conditional proof. The role of a conditional proof is to *prove* that a certain conditional follows from specific premises that are not themselves assumptions (and that will not be discharged). For instance, the conditional proof we considered above is a proof that $A \supset C$ follows from $A \supset B$ and $B \supset C$. This proof just serves to establish that transitivity of implication holds for material conditionals—that if we know that both $A \supset B$ and $B \supset C$, we will always acquire a true belief in judging that $A \supset C$. An inferential procedure that takes the first two material conditionals as premises and outputs the third as a conclusion will be truth-preserving. This suggests an obvious alternative for understanding the transitions in psychological states that actually occur when we infer that $A \supset C$ from $A \supset B$ and $B \supset C$. Supposing we believe that

A ⊃ B and B ⊃ C, we may infer directly from those beliefs that A ⊃ C. The inference rule followed would be transitivity of implication. So, to explain our ability to come to know A ⊃ C on the basis of knowing A ⊃ B and B ⊃ C, we need only posit an ability to reason in conformity with transitivity of implication.

Of course, it remains *possible* that we might, instead, break the inference into additional steps "in imagination"—steps mirroring the steps of a conditional proof—representing that A, B, and C via *sui generis* imaginative states, before concluding that A ⊃ C. This would be, in effect, an alternative method for carrying out the same computation of deriving A ⊃ C from the inputs of A ⊃ B and B ⊃ C. This latter method modelled on the method of conditional truth has the virtue of not requiring use of the inference rule of transitivity of implication. Yet it has the vices of both requiring additional inferential steps and the interaction of two different kinds of mental states (beliefs and *sui generis* imaginings). If we really did, at the psychological level, carry out an inference mirroring the steps of this conditional proof, we would need to keep in mind five premises at once—two of which involve conditionals—in order to arrive at the conclusion. The alternative method, which moves directly from two premises to the conclusion, has the virtue of requiring fewer steps and of only employing beliefs. It has the corresponding vice of requiring use of an extra inferential rule: transitivity of implication. From my vantage, this method appears simpler overall; from any vantage, it is at best a toss-up. This case of conditional proof gives no reason to posit *sui generis* imaginative states.

Stepping back, it's easy to see that any case of conditional proof at all can be reconceived without assumptions—even while still working within a framework that otherwise mirrors closely the steps of a proof within a system of natural deduction. The method of conditional proof simply serves to show that it is truth-preserving to infer a certain conditional from a certain set of (not-to-be-dischared) premises. Doing without the assumptions requires exploiting an additional inferential rule in their place—one that takes us from the main premises to the conclusion. Granted, it may seem extravagant to posit a new rule for each species of inference (to a conditional) we might wish to make. But, as a practical matter, most of us will be unable to conduct many different species of such proofs "in the head" anyway—whether we think of them as involving assumptions or not! If our inferential capacities are limited, in practice, we needn't attribute to ourselves a grasp of all the inferential rules we might need, in principle.

5.7.2 *Reductio* without Assumptions

Similar points apply to the case of *reductio ad absurdum* in systems of natural deduction, where we "assume" that *p* as a means to establishing that *not-p*. When we carry out a *reductio* "in our heads"—assuming that *p* in the process—it may seem that we exploit a mental state *other than* one of our beliefs to record the assumption. Here again we seem to have reason to posit *sui generis* imaginative states. Perhaps we assume that *p* by tokening an imaginative state with the content *p* and then appreciate the contradiction that follows "in imagination."

To see why psychologizing this procedure needn't involve *sui generis* imaginative states after all, it will help to consider a couple of concrete examples. I offered one style of *reductio* above in passing, when dismissing the thesis that subjunctive (counterfactual) conditionals have the same truth conditions as the material conditional. We know that the material conditional is true whenever its antecedent is false. Therefore, if counterfactual conditionals have the same truth conditions as the material conditional, then all counterfactual conditionals with false antecedents are true. But that is absurd, because it is clear that many counterfactuals with false antecedents are false (e.g., "If I had dropped a feather on my toe, it would have left a bruise"). Therefore, we are warranted in rejecting the claim that counterfactual conditionals have the same truth conditions as the material conditional. Writing the proof out in steps, it might look like this:

1. Counterfactual conditionals have the same truth conditions as the material conditional. (Assumed for *reductio*)
2. The truth conditions for the material conditional mandate that a material conditional is true whenever its antecedent is false. (Premise)
3. All conditionals with the same truth conditions as the material conditional and with a false antecedent will be true. (Lemma, from 1, 2)
4. Therefore, all counterfactuals with false antecedents are true. (Lemma)
5. Not all counterfactuals with false antecedents are true. (Premise)
6. Steps 4 and 5 generate a contradiction; reject 1, 2, or 5.

If we were to psychologize these steps as a means to understanding how the computation is carried out psychologically, steps 1 and 4 would correspond to *sui generis* imaginative states; for we can assume that the person carrying out the inference does not belief those propositions. (Those propositions are not "in" their Belief Box.) The other steps, presumably, correspond to beliefs in one's knowledge store.

The purpose of a *reductio* is to arrive at an answer as to whether to a particular proposition (to be rejected) is true. Put in terms of an input and output, and of a process that mediates between them, the input can be seen as a question, namely: is it the case that *p*? The output of a successful *reductio* returns the answer: "No." Staying with the example above, let us suppose that the process begins by the subject registering the question "Is it true that counterfactual conditionals have the same truth conditions as the material conditional?" The system searches its knowledge store for information relevant to answering and, in particular, locates 1, 2, and 3:

1. All material conditionals with false antecedents are true, as a matter of their truth conditions. (Premise)
2. If counterfactual conditionals have the same truth conditions as the material conditional, then all counterfactual conditionals with false antecedents are true. (Premise)
3. Not all counterfactuals with false antecedents are true. (Premise)
4. Therefore, counterfactual conditionals do not have the same truth conditions as material conditionals. (Conclusion, from 2, 3 *modus tollens*)

This reasoning process answers the same question as the one modelled on *reductio ad absurdum*. It takes, as input, the question of whether counterfactual conditionals have the same truth conditions as the material conditional and gives, as output, the answer: No. And it makes use of essentially the same stored information. However, it does not call upon any internal states that are not beliefs—no "merely assumed" representations. Nor does it exploit any unusual rules of inference. For all we know a priori, when humans carry out the kind of reasoning associated with this style of *reductio*, they are, at the psychological level, making inferences from among their beliefs, in the manner of 1–4, using *modus tollens*.[7]

This style of *reductio*, while common, is not a *reductio* in the strict sense. It relies upon a strong prior belief (in 3) that conflicts with an entailment of the proposition in question. More formal arguments by *reductio* work simply by showing how a contradiction follows from a certain premise—a premise that is then rejected because of its entailment of a contradiction. Such arguments do not rely upon one's having a prior conviction that the denial of a certain proposition (e.g., step 3) would be unacceptable. These might seem more clearly to necessitate *sui generis* imaginative states as such cases don't lend themselves to reformulation as instances of *modus tollens*. Let us consider an example of this kind, which I adopt from Rescher (2018). Suppose that we are uncertain whether it is possible to divide a non-zero number by zero to get a well-defined quantity, *Q*. A classical *reductio* of

[7] I hear the following objection: "Let A be 'Counterfactual conditionals have the same truth conditions as material conditionals' and let B be 'All counterfactual conditionals are true.' It is precisely by representing that A, in imagination, that we come to infer, in imagination, that B! And it is again only by representing that B, in imagination, that we are able to see that it conflicts with our belief that not-B."

The heart of the objection is that it is only by representing that A, in imagination, that we are able to *dwell upon* A so as to see that B follows from it; and then it is only by representing that B, in imagination, that we are able to *dwell upon* it so as to see that it (absurdly) conflicts with our stronger prior belief that not-B. But this objection leads nowhere. Imagination is not required to dwell upon questions generally. When asked, we may dwell upon the question of our favorite pizza topping. Turning our attention to that question does not require us to enter into any *sui generis* imaginative states. The same goes for turning our attention to the question of what would happen if A, or the question of whether anything we believe conflicts with B. (I expand on this point in section 8.8.) Whether *answering* the question we have so posed requires *sui generis* imaginative states is the better, more difficult question I am considering at length.

the proposition that this is possible would begin with the assumption whose truth we wish to assess:

(1) $x \neq 0$ and $x \div 0 = Q$ (Assumption)

By familiar principles linking multiplication and division, we can then derive (2):

(2) $x = Q \times 0$
(3) Any number multiplied by 0 is 0. (Premise)
(4) Therefore, $x = 0$.

Step (4) contradicts our assumption in (1) that $x \neq 0$. It is the assumption of (1) itself, together with bedrock principles of arithmetic, that leads to a contradiction of (1) with (4). Noticing the contradiction, we can either reject the bedrock principles of arithmetic or reject (1) on the grounds that its truth entails a contradiction. As written, steps (1), (2), and (4) are all suggestive of states that are not believed and would need to be "merely imagined." Our question is whether this same computation can be carried out through a belief-only reasoning process. We need to start by clarifying the nature of the reasoning: it takes, as input, the question: When $x \neq 0$, and x is divided by 0, can x be Q? As output, it gives the answer: No. On reflection, we can see how the relevant reasoning could instead exploit transitivity of implication, discussed above on the topic of conditional proof. We can rely upon background knowledge of arithmetic principles to infer as follows:[8]

Revised reductio

(1) If $x \neq 0$, and $x \div 0 = Q$, then $x = Q \times 0$

This inference is made on the basis of the same knowledge that allows us to have (2) as a premise in the *reductio*, as initially written. Then, recalling the principle that 0 multiplied by any number is 0, we can infer:

(2) If $x = Q \times 0$, then $x = 0$.

Finally, by transitivity of implication for the material conditional, we are able to conclude:

(3) If $x \neq 0$ and $x \div 0 = Q$, then $x = 0$. (1,2 by transitivity of implication)

When we arrive at (3), we see that something has gone wrong. Our conclusion has the form: if not-A & B, then A. This conditional is not itself a contradiction. By the logic of the material conditional, (3) will be true whenever "$x \div 0 = Q$" is

[8] Whether we actually do so will of course depend upon our facility with mathematics—as it will no matter how we understand the algorithm. The point is to show that there is no practical computational limit imposed by working only with beliefs.

false—i.e., it will always be (vacuously) true. But recall the question that started the computation: when $x \neq 0$, and x is divided by 0, can x be Q? We have found, in (3), that if $x \neq 0$ and $x \div 0 = Q$, then $x = 0$. It is now clear that the truth of the antecedent implies its own falsity. Or so we believe, if we believe (3). This is reason to reject the truth of the antecedent itself—just as step (4)'s contradicting step (1) in *Classic Reductio* gives reason to reject (1). We have our answer to the question that began the computation: No, it can't be that $x \neq 0$, and $x \div 0 = Q$.[9] The same function computed in *Classic Reductio* has been computed without the use of *sui generis* imaginative states, while limiting ourselves to the tools of natural deduction itself. Further, doing so hasn't forced us into computations of obviously greater complexity.

With these examples, I don't claim to have established that every conceivable *reductio* could receive this kind of treatment. But then, we already knew, from reflection on axiomatic proof systems, that it is *possible* to make do without assumptions within a formal system for deduction. What I have aimed to show, in this section, is that *even when working within the general terms of a system of natural deduction*, the denier of *sui generis* imaginative states has ample room to maneuver when faced with explaining deductive inferential patterns that appear to require something like a cognitive state of "assuming," "supposing," or "imagining." The apparent need for "assumptions," even within those systems, is *only* apparent, if we allow ourselves a few tweaks. We have also seen that systems of natural deduction are themselves limited in their capacity to model the actual inferential patterns of human thought—including, especially, thought about conditionals and hypothetical entailment. The combined upshot is that the occurrence of assumptions within the steps of proofs in systems of natural deduction gives us little reason to posit corresponding *sui generis* imaginative states.

5.8 Mental Models?

I noted earlier that a number of psychologists advocate a "mental models" approach to the psychology of conditional reasoning, developed most prominently by Philp Johnson-Laird and colleagues (Byrne, 2005; Johnson-Laird, 1983; Johnson-Laird & Byrne, 2002). The authors who posit mental models sometimes describe the use of such states as "imagining," and hold that they are exploited when people consider whether various forms of reasoning are deductively valid.

[9] We get a similar result when interpreting the conditional in (3) in line with the Ramsey test, discussed next chapter. On that view of conditionals, (3) should only be believed if "$x = 0$" has a high probability within a belief set containing "$x \neq 0$" and "$x \div 0 = Q$." Obviously, "$x = 0$" will not receive a high degree of probability within any belief set containing "$x \neq 0$," and so (3) will be rejected. We will not, in turn, believe the antecedent of (1)—"x is not 0, and $x \div 0 = Q$"—because we see that any belief set that contains it must contain (3) as well, which has been rejected.

So it is worth exploring whether these theorists are in fact committed to the kind of *sui generis* imaginative states that I've claimed we needn't countenance. The reader is forewarned that translating the terms of psychologists into those familiar to philosophers is not always straightforward. Before beginning, let me preview my conclusion: psychologists positing mental models often hold that the models consist in sequences of mental imagery; they therefore assume that uses of mental models are cases of I-imagining, in my sense. However, when we consider the attitude or force of those imagistic model-states, the most natural interpretation is that they are judgments or beliefs. Specifically, the mental models are *proper parts of* judgments or beliefs. Thus, these theorists are not committed to anything like a *sui generis attitude* of imagining. Their views are compatible with reducing A-imagining to a collection of more basic folk psychological states.

Johnson-Laird and Byrne ("JLB") see their theory—which extends beyond reasoning with conditionals to include inductive and deductive inference generally—as differing from what they call "formal rule theories." Formal rule theories, by their reckoning, are classical computational approaches according to which "individuals reason using formal rules of inference like those of a logical calculus" (2002, p. 646). They propose that instead of using quasi-logical formal rules of inference, reasoners "imagine the possibilities under consideration—that is, [they] construct mental models of them" (p. 647). The "underlying deductive machinery" at work in conditional reasoning, they argue, "depends not on syntactic processes that use formal rules but on semantic procedures that manipulate mental models" (p. 647). They support their theory with experiments showing that people reason about conditionals in ways that would be expected if they were using mental models of a specific sort (and not "formal rules").

How do mental models relate to things philosophers are accustomed to theorizing about—such as propositional imaginings, or sensory imaginings? JLB's broad characterization of mental models doesn't clarify matters. Mental models, they propose:

> can be constructed from perception, imagination, or the comprehension of discourse. They underlie visual images, but they can also be abstract, representing situations that cannot be visualized. Each mental model represents a possibility. It is akin to a diagram in that its structure is analogous to the structure of the situation that it represents, unlike, say, the structure of logical forms used in formal rule theories. (2002, p. 647)

On the one hand, mental models can "underlie visual images" and so, perhaps, are imagistic in nature.[10] This seems to fit with their idea that, unlike "logical

[10] Elsewhere: A model "may take the form of a visual image" (Johnson-Laird, Byrne, & Schaeken, 1992, p. 421). "The end product of perception is a model of the world (Marr, 1982)" (1991, p. 421).

forms," a model's structure "is analogous to the structure of the situation that it represents." On the other hand, these models can also be "abstract, representing situations that cannot be visualized." In that case, it is unclear how we are to understand the structural isomorphism between the representation and its content. JLB invoke imagination (as well as perception, and language comprehension) as a kind of faculty that *constructs* mental models; but they say little else about the faculty. More often, when they speak of imagining, they appear simply to have in mind the occurrent use of mental models—whatever *their* nature.

We are able to get a clearer picture of the relation between mental models and folk psychological states like belief by looking at the role mental models play in JLB's theory. "By definition," they tell us, "a mental model of an assertion represents a possibility given the truth of the assertion. Hence, a set of mental models represents a set of possibilities" (p. 653). Of course, actualities are possibilities; thus, when representing actualities, beliefs represent possibilities. So, the fact that a mental model represents a possibility is not at odds with its being a constituent of a belief. In the case of what they call "basic conditionals,"[11] JLB hold that there are three possibilities where the conditional "if *a* then *b*" is true, mirroring the three rows of the truth table for the material conditional where a conditional is true.[12] In one such possibility, *a* and *b* are both true; in another, *a* is false while *b* is true; in the third, both *a* and *b* are false. To represent that set of three possibilities, JLB propose, one needs to generate three separate mental models: one mental model representing *a* and *b* as both being the case; another representing not-*a* and *b*; and a third representing not-*a* and not-*b*. JLB use quasi logical notation in symbolizing these "models" as follows:

a b
not-*a* *b*
not-*a* not-*b*

Anyone who explicitly represents what they call the "core meaning" of a basic conditional will generate all three mental models simultaneously, on their view. However, they claim, reasoners typically do not *explicitly* represent all three mental models when considering a conditional. Rather, they often represent some of the possibilities only "implicitly"—in particular, those where the antecedent is

[11] Basic conditionals, for Johnson-Laird and Byrne, "are those with a neutral content that is as independent as possible from context and background knowledge, and which have an antecedent and consequent that are semantically independent apart from their occurrence in the same conditional" (2002, p. 648).

[12] In contrast to Evans & Over (2004), who posit a role for mental models in conditional reasoning while espousing the Ramsey test (see Chapter 6) as a criterion for a conditional's acceptability, JLB hold that the truth conditions for conditionals are in fact those of the (truth-functional) material conditional—even if we do not use "formal rules" when assessing them. See Barrouillet et al. (2008) for further discussion.

false. To represent the model implicitly is to be *disposed* to generate the model explicitly, should one be triggered in the right way.[13] On their theory, very often the only model explicitly represented (read "explicitly represented" as *occurrently tokened*) when someone thinks about a conditional is the first of the three mentioned above, where the antecedent and consequent both hold. They symbolize this tendency—where one model is explicitly represented, with others represented only implicitly—by showing a model of the first situation (true antecedent and consequent), with an ellipsis below it:

a b
…

The ellipsis symbolizes that one is disposed to generate certain other models, but is not yet doing so. JLB argue that our tendency to generate just one of the three models in the set corresponding to a conditional (with the others "footnoted") serves to explain various experimental results on conditional reasoning. One is that *modus ponens* is an easier inference form to process than *modus tollens*; a second is that people are more likely to fall into the error of affirming the consequent than denying the antecedent (pp. 666–9).

Whether their theory is fact well supported by such findings needn't concern us here. Our interest is in the nature of mental models themselves—in whether we would need to countenance something like *sui generis* imaginative states if we wished to avail ourselves of JLB's theory. We now have enough pieces of the theory on the table to see that we do not. For it is not only conditionals that are represented via mental models on their theory; practically any other kind of assertion is as well. Consider the inclusive disjunction: either *p*, or *q*, or *p* & *q*. According to JLB, a person who explicitly represents (and occurrently judges) such a disjunction to be true forms the following set of mental models:

p
q
p q (p. 653)

Three distinct mental models are tokened as a means to representing the single proposition that *p*, or *q*, or *p* & *q*. Do we *imagine* these three "possibilities" when we explicitly represent the inclusive disjunction? Perhaps we do so in the

[13] JLB explain this as follows: "Basic conditionals have mental models representing the possibilities in which their antecedents are satisfied, but only implicit mental models for the possibilities in which their antecedents are not satisfied. A mental footnote on the implicit model can be used to make fully explicit models… but individuals are liable to forget the footnote and even to forget the implicit model itself" (2002, p. 654).

imagistic sense of imagine, supposing we use mental imagery in the process. But the question we are interested in is whether we generate a *sui generis* imaginative state in doing so—some state that cannot be viewed as a belief. (We've already seen, in Chapters 3 and 4, that there is no barrier to beliefs having mental images as proper parts.) Here the answer must be no. What JLB have given us is an account of what it is to believe an inclusive disjunction—or, perhaps better, to *occurrently* believe or *judge* an inclusive disjunction to be true. Namely, it is to think about these three possibilities simultaneously, via the use of this set of three mental models. To judge the proposition that *p, or q, or p &q* is, for JLB, nothing other than to token these three mental models. Mental models appear to be the constituents of beliefs, then, and not *sui generis* states that stand apart from them. Indeed, for JLB, the only difference between representing "if A then C" and the conjunction "A and C" lies in what is *implicitly* represented—i.e., the models we are *disposed* to generate, when triggered in the right way. In explaining the role of the ellipsis in their account, they note that "the ellipsis denotes the implicit model, which has no explicit content, and which distinguishes a conditional from a conjunction, *A and C*" (p. 655). So, on their account, there is often no *explicit* cognitive difference in what is represented when one represents a conditional and when one represents a simple conjunction. In both cases we explicitly "imagine" the same possibility where *a* and *c* hold, the only difference lying in certain "implicit" models (with "no explicit content") being available—if they are triggered in the right way—in the case of the conditional. This helps to clarify that simply judging that A and C often involves the same explicitly represented mental models as judging that if A then C; there is obviously no clash here with mental models serving to realize beliefs.

We can, when prompted, go on to explicitly represent the other two (normally implicit) models that differentiate believing the conditional from believing the conjunction. But this still amounts to believing (now *completely explicitly*) the conditional. We still have not in any sense stepped outside of what we really believe. Thus, JLB do not have anything like the notion of a cognitive attitude of imagination in mind when they speak of "imagining possibilities." They are better seen as making a claim about the nature of judgments. They are saying that *what it is* to judge "if *p* then *q*" is to generate one or more of the three mental models listed above (and that, typically, we generate just one, which explains various experimental results). There is no notion of a *sui generis* imaginative state at work. They do suggest that these models are often imagistic in nature. But, again, there is no tension in the idea that some beliefs and judgments have mental images as constituents. So, despite appearances—and despite JLB's own affinity for characterizing uses of mental models as "imagining possibilities"—there is nothing in their theory of mental models and conditional reasoning that stands in the way of explaining conditional reasoning entirely in terms of sequences of (sometimes imagistic) beliefs.

5.9 Summary

Let's recap. I distinguished three kinds of conditional—material, indicative, and subjunctive—and explained, briefly, why most take the conditionals of natural language (indicatives and subjunctives) to behave differently than the material conditional. Nevertheless, we may at times reason in accord with the material conditional. Philosophers, especially, are often inclined to characterize human thought in such terms. So it is worth considering whether systems of natural deduction, in which the material conditional occurs, give us reason to think that imagination cannot be explanatorily reduced to a collection of more basic folk psychological states. The best reasons appear to lie in two species of deductive reasoning that involve "assuming true" a proposition one does not believe: the method of conditional proof, and arguments via *reductio*.

In response, I first noted that the existence of axiomatic proof systems shows that there is no *in principle* difficulty in doing without "assumptions" (or corresponding *sui generis* imaginative states). Second, I discussed empirical work that casts strong doubt upon the claim that human reasoning mirrors the steps of a proof in a system of natural deduction. Even when contextual features are removed from reasoning tasks, so as to highlight their abstract structure, ordinary participants do not evaluate conditionals as they would if the logic of the material conditional were mirrored in their inferential architectures. This robust finding has led psychologists to seek means other than systems of natural deduction for modelling the psychological processes at work in "abstract" conditional reasoning tasks, such as judging the validity of a pattern of inference, or assessing the truth conditions of an artificially concocted conditional (as in the Wason selection task). Third, looking more closely at a few deductive proofs that make paradigmatic use of assumptions, I showed that the same conclusions can be reached without the use of assumptions, while still limiting oneself to a framework similar to that of natural deduction. This helps to further chip away at the sense that assumptions—conceived of as a *sui generis* mental states akin to "imaginings"—are especially valuable theoretical posits. The discussion concluded with a close look at the notion of a "mental model" as it appears in the influential work of Johnson-Laird and Byrne. Mental models, I argued, are best viewed as constituents of occurrent judgments—and, possibly, desires as well—and not as *sui generis* imaginative states.

Having seen that the material conditional, as it functions within systems of natural deduction, does not give us reason to posit *sui generis* imaginative states, we can now turn to see how matters stand with respect to the indicative and subjunctive conditionals of natural language.