What Reasoning Might Be

The philosophical literature on reasoning is dominated by the assumption that reasoning is essentially a matter of following rules. This paper challenges this view, by arguing that it misrepresents the nature of reasoning as a personal-level activity. Reasoning must reflect the reasoner's take on her evidence. The rule-following model seems ill-suited to accommodate this fact. Accordingly, this suggests replacing the rule-following model with a radically different, semantic approach to reasoning.

1. Introduction

Reasoning is an activity familiar to all of us. But what exactly does one do when one reasons? For example, consider a subject who knows the following:

1. If Socrates is human, then he is mortal
2. Socrates is human

We naturally think that there is a cognitive act — albeit a rather trivial one, in this particular example — that the subject can perform in order to get to know the following:

3. Socrates is mortal

What is the nature of this cognitive act?

According to many, reasoning is fundamentally a matter of following rules.¹ This is not just the relatively innocuous claim that reasoning (or at least good reasoning) can be described or captured by rules. It is the stronger claim that our subject gets to know that Socrates is mortal in

¹ What does it mean to say that reasoning, or any other activity, fundamentally consists in Φ-ing? Consider what the activity of playing basketball consists in. One might answer this question on many different levels, including the anatomic/physiological level, the level of individual movements, and the level of strategy and tactics. But there is a sense in which more fundamental than all of those is an abstract specification of what the game is all about: roughly, two teams competing against each other, scoring points by getting the ball through hoops mounted on poles. This level of description is fundamental in the sense that descriptions at all other levels are intelligible by reference to this one: they are further specifications of how one does what is specified at this level of description. This is the sense in which, according to the rule-following model, reasoning is fundamentally a matter of following rules.
virtue of being guided by, or following, a rule — a mental analogue of the rule of modus ponens familiar from propositional logic.

The rule-following model is not often explicitly defended against alternatives; much contemporary work on the nature of reasoning simply appears to take it for granted (for important exceptions see Boghossian 2014 and Broome 2013; 2014). This should be surprising, because (as we shall see) the rule-following model is at odds with some deep-seated intuitions about the nature of reasoning. Reasoning, as a personal-level activity, seems to be a paradigm case of the kind of control that we have over our own minds: it is our making up our own minds. One natural corollary of this idea is that reasoning must reflect the subject’s own take on her evidence. The rule-following model has trouble accommodating this thought.

The fact that the rule-following model faces trouble in this area has not gone unnoticed. Paul Boghossian (2003; 2008; 2014), in particular, has written extensively and forcefully on the topic. Recognizing those difficulties, however, has not led Boghossian to reject the rule-following model: on the contrary, he suggests accepting rule-following as a basic and unanalyzable mental capacity (Boghossian 2014, 16-18). Boghossian, like other proponents of the rule-following model, does not even consider alternatives. My positive goal in this paper is to develop just such an alternative.

On the rule-following model, the rules that guide reasoning are formal, in a sense analogous to the sense in which the rules of inference that characterize a formal system are: just as the latter deal only with syntactic objects within the system rather than with the subject-matter the system is intended to capture, the rules of reasoning deal only with our attitudes, rather than

---

2 It is impossible to give an exhaustive list of philosophers who endorse the rule-following model of reasoning, since so often the model is simply take for granted. For a sampling of authors — in addition to Boghossian and Broome — who are more or less explicit on the issue, see Winters (1983), Van Cleve (1984), Brandom (1998), Wedgwood (2002), Wright (2004a; 2014), Hlobil (2014), and Miller (2015). This is not to say that the rule-following model has gone entirely unchallenged in philosophy. Ian Rumfitt (2008; 2011), for one, proposes an alternative that is in many ways similar to my own.
the subject-matter of our reasoning. Intuitively, however, our reasoning is not guided by thoughts about our attitudes and their contents; it is guided by thoughts about the world. This suggests that we should think of reasoning in semantic terms. As I hope to show, such an approach to reasoning not only makes it easy to accommodate the role of the reasoner’s own take on her evidence in the activity of reasoning, it also allows for a more satisfying account of the place of reasoning in our cognitive lives.3

Finally, a word about the scope of this paper. Reasoning, understood broadly as the activity of bringing our attitudes in line with the norms of reason, is encountered in the wild in a variety of guises: argumentation, problem-solving, decision-making, planning, theory-building and testing all involve reasoning. Despite this variety, my attention in this paper will be almost exclusively on theoretical reasoning — and especially simple cases of theoretical reasoning at that. This is not because I believe that all forms of reasoning ultimately reduce to reasoning of this type. I am open to the possibility that other forms of reasoning might involve genuinely new features, beyond anything discussed here. It is, rather, because of the more modest hope that the relative simplicity of the cases I discuss will allow us to gain insights that will then be applicable to the more general case as well. In particular, if — as I shall argue — the rule-following model fails even in this simple case, it seems likely that we will need to reconsider its applicability across the board.

3 The “mental models” theory (Johnson-Laird 1983; 2001; 2008; Johnson-Laird and Byrne 1991) in the psychology of reasoning is also often advertised as “semantic”, and is specifically developed in reaction to rule-based accounts such as Rips 1994. However, care is needed in interpreting this claim (I will briefly return to this in Section 5). For now, I just want to note that my aim is to answer a rather different question from the one that mental model theory aims to answer: my concern is what you do when you reason, rather than how reasoning is carried out at the computational level. Of course, the two questions are not simply independent of each other: an account of what we do when we reason must be sensitive to much of the same empirical data as an account of how reasoning is carried out at the computational level, while the latter sort of account can benefit from a clearer conceptual characterization of the phenomenon it seeks to explain.
2. The Taking Condition

My argument relies on a certain condition upon theories of reasoning — namely, that they must explain how reasoning reflects the subject’s own take on her evidence. My aim in this section is to motivate and explain this condition.

Paul Boghossian (2014, 4) quotes approvingly the following characterization of inference, by Frege (1979, 3):

To make a judgment because we are cognisant of other truths as providing a justification for it is known as inferring.

As Boghossian notes, Frege’s characterization has to be amended to allow for inferences based on false premisses, as well as for inferences in which the premisses do not actually support the conclusion. But for present purposes the central feature of Frege’s characterization is the claim that inferring \( p \) from a set of premisses \( R \) requires taking \( R \) to provide justification or support for \( p \), and coming to believe \( p \) (partly) because of this. Following Boghossian (2014, 5), I will call this the “Taking Condition” on inference. Is it a reasonable condition?

Some authors use the terms “reasoning” and “inference” not just for personal-level performances, but also for sub-personal information processing. For instance, humans are pretty good at judging the emotions of other people on the basis of subtle facial and behavioral cues. Some authors would be happy to take such judgments to be the conclusions of unconscious inferences (e.g., Johnson-Laird 2008, 60–72). On such broad usage the Taking Condition seems clearly false: when making such a judgment you need not be aware of the grounds on which you have made it. Thus, in endorsing the Taking Condition I am implying that such judgments are not inferences. But why should the application of the terms “reasoning” and “inference” be restricted in this way?

The reason is that such broad usage obscures a crucial point. In one central sense of these terms, reasoning or inferring are things that we do. Reasoning is an expression of the sort of control that we have over our cognitive lives; it is our way of making up our own minds. One way to
bring this fact out is by noting that it has distinctive normative import: if you make a bad inference, we can legitimately criticize you as having been hasty, irresponsible, biased, and so on. By contrast, there is only a very thin sense in which the subject herself is responsible for her immediate judgments about another’s emotional state. If you misread another’s facial expressions, your mistake is more akin to a perceptual illusion than a case of bad reasoning. A natural way to explain this difference is to say that inferring reflects the subject’s take on what her evidence requires, whereas our system for judging other people’s emotional states is a lower-level system whose workings are opaque to us. Marking this difference is the point of the Taking Condition.⁴

Is the Taking Condition sufficient to capture the full sense in which reasoning is something that we do? I will consider this question in more detail in Section 5, once my positive account of how we meet the Taking Condition is on the table. For now, it is worth noting that nothing in my argument depends on whether the Taking Condition is a sufficient condition for reasoning; it is enough if it is a necessary one.

But is the Taking Condition even a necessary condition for reasoning? John Broome has recently argued that it is not (Broome 2014, 19–20). The problem is that, according to Broome, the Taking Condition does not generalize to practical — and in particular, instrumental — reasoning. The reason is supposed to be that while relations of support among propositions are

---

⁴ An anonymous referee asks how this distinction relates to the “System 1/System 2” distinction, made famous by Daniel Kahneman (2011). I think there is no simple answer to this, as Kahneman’s System 1 is a very mixed bag — bundling together things like the processing that underlies depth perception with “intuitive” responses to logical or mathematical puzzles (Kahneman himself, of course, recognizes that he abuses terminology in his labeling). But while depth perception is quite plausibly handled by an informationally encapsulated and cognitively impenetrable system, our intuitive responses to, e.g., the Linda the bank-teller case are not — witness the fact that, once you see the right answer it no longer seems to you that the wrong answer is correct (though, strikingly, this does not seem to immunize you from similar mistakes in the future). This sort of difference suggests that we should not expect a theory of reasoning to treat both of these processes on a par.
relevant to theoretical reasoning, they are irrelevant to practical reasoning. For example, the proposition that Mary is making an omelette arguably supports (in the sense Broome has in mind) the proposition that she will break some eggs; but this fact seems entirely irrelevant to the way in which Mary herself forms the intention to break some eggs. From Mary’s point of view, what matters is the fact that her intention to make an omelette rationalizes an intention to break some eggs, rather than a relation between the bare contents of those intentions.

But I think Broome’s argument rests on a rather implausible understanding of the relation of support involved in the Taking Condition. As I understand this relation, saying that $R$ supports $p$ means that committing to the truth of all members of $R$ rationally commits you to the truth of $p$ as well. Alternatively, it means that, given your commitment to $R$, all ways for things to be that are epistemically possible for you are such as to make $p$ true as well. Thus support is not just a bare relation among propositions; it is a relation among commitments to propositions. If we understand the relation of support in this way Broome’s objection lapses — for it is very plausible that Mary’s instrumental reasoning does involve her recognizing that her commitment to omelette-making also commits her to egg-breaking. (Alternatively, it involves her recognizing that, consistent with her commitment to making an omelette, the only options practically open to her involve egg-breaking.) No doubt more needs to be said to clarify the nature of the relation of support relevant to the Taking Condition. I will say a bit more, concerning the theoretical case specifically, in Section 4; for present purposes what matters is that there seems to be no reason to worry about the generality of the Taking Condition.

Another question that may be raised concerns the type of state the takings required by the Taking Condition exemplify. I would like to leave this question as open as possible. Consider the following case. Tom has some unjustified or irrational theoretical beliefs. For example, he believes that certain spots on people’s faces indicate that they have been marked by a demon, and once so marked they will soon die. Tom sees such spots on Bob’s face. As a result of his theoretical beliefs, he takes it that the spots on Bob’s face support the conclusion that Bob will
soon die. Thus he reasons from the premiss that Bob has spots of such-and-such a sort on his face to the conclusion that he will soon die. As it happens, the spots on Bob’s face are a sign of advanced disease, and so their presence does in fact indicate that Bob will soon die — though for reasons Tom has no inkling of. Is Tom’s reasoning rational in this case?

I am inclined to say that Tom’s reasoning is not rational. Intuitively, the belief that results from Tom’s reasoning — his belief that Bob will soon die — is not epistemically justified. But this belief is, by hypothesis, derived just from Tom’s belief that Bob has such-and-such a pattern of spots on his face; and this belief is justified. Barring some unappealing restriction of closure for epistemic justification, therefore, it looks like we have to conclude that there is something wrong with the way in which Tom reasons. I think that the flaw is not hard to find: although the presence of the spots on Bob’s face does support the conclusion that he will soon die, Tom (in light of his irrational theoretical beliefs) is not justified in taking it to support this conclusion. This implies that the takings required by the Taking Condition must exemplify states that can be assessed for rationality or epistemic justification. Since belief is the paradigm case of such a state, I am happy to think of them as beliefs.

This, however, conflicts with views such as Tucker’s (2012) and Broome’s (2013; 2014) that rely on intellectual seemings. Since seemings are not the sort of thing that can be either rational or irrational, such views would have trouble finding anything wrong with Tom’s inference above. (This is a bullet Broome and Tucker may be happy to bite.) As it happens, nothing in what follows hinges on taking a stand on this particular question. Thus, while I will sometimes use doxastic language in speaking of the takings required by the Taking Condition, it should be kept in mind that something less than belief — something, in particular, not assessable for epistemic justification — might suit as well.

---

5 That is, Tom reasons directly from the premiss that the relevant pattern of spots is present on Bob’s face to the conclusion that Bob will soon die. Tom’s crazy theoretical beliefs are part of the background of his inference, not among his premises. Tucker (2010) questions this distinction, but I think it is essential for understanding inference.
Let us now turn to the question whether rule-following theories of inference can explain how reasoners meet the Taking Condition — or, alternatively, whether they provide a different way of capturing the desiderata that the Taking Condition is meant to capture.

3. The Taking Condition and Rule-Following Theories of Inference

Consider a subject performing the elementary inference from Section 1:

1. If Socrates is human, then he is mortal
2. Socrates is human
3. Therefore, Socrates is mortal

Assuming that the Taking Condition holds, our subject must take it that (1) and (2) support (3), and come to believe that Socrates is mortal partly because of this. Can rule-following theories explain this?

Let us begin by considering what the rule governing this bit of reasoning might be. It is not the familiar modus ponens rule of propositional logic, as that is a rule about strings of symbols in a formal language, not a rule of reasoning. We want something analogous, but concerning transitions among beliefs. Such a rule might be formulated as follows:

(MP) If you are rationally permitted to believe both that \( p \) and that ‘If \( p \), then \( q \)’, then you are prima facie rationally permitted to believe that \( q \). (Boghossian 2008, 472)

Suppose that our subject is reflective and logically astute, and so can plausibly be said to believe that (MP) is a good rule. Can this belief help explain how our subject meets the Taking Condition? It clearly cannot do so directly. After all, (MP) says nothing about Socrates or his mortality; so how can it explain our subject’s coming to believe that Socrates is mortal?

One way for it to do so would be this. Our subject can substitute in (MP) as follows:6

---

6 This way of presenting matters presupposes that rules of reasoning are schemas of truth-evaluable statements, so that their substitution-instances can figure as premisses in reasoning. But this is not
a. If I am rationally permitted to believe that Socrates is human, and that if Socrates is human then he is mortal, then I am prima facie rationally permitted to believe that Socrates is mortal.

Assume, further, that our subject knows what she believes, and moreover that she is justified in taking her own beliefs to be rationally permissible ones (although neither of these assumptions is entirely innocent, of course). Then she can also rely on the following premiss:

b. I am rationally permitted to believe that Socrates is human and that if Socrates is human then he is mortal.

And from these two premisses, she can conclude:

c. I am rationally permitted to believe that Socrates is mortal.

Assuming, now, that coming to believe (c) is sufficient to get our subject to form the first-order belief that Socrates is mortal, her reasoning is done. Familiarly, however, nothing like this will do as a fundamental account of how we reason. One way to see this is simply to note that the transition from (a) and (b) to (c) in this argument is itself a modus ponens transition. Taking the subject to perform this inference, therefore, presupposes exactly the capacity that we were hoping to explain.  

What the above argument highlights is that we need to avoid explaining the sense in which a subject is guided by her grasp of a rule in inferential terms: if you need to reason in order to follow a rule then rule-following cannot be what you fundamentally do when you reason. The natural alternative is a dispositional account of rule-following: your doing \( x \) in circumstances \( C \) .

---

7 The history of arguments in this vicinity traces back at least to Lewis Carroll’s (1895) story of Achilles and the Tortoise. Variations are given by Winters (1983), Van Cleve (1984), Johnston (1988), Brewer (1995), Fumerton (1995), Boghossian (2003; 2008; 2014), Raiton (2006), Broome (2013), Wedgwood (2006), and others. For some replies, see Leite (2008) and Valaris (2014). Worries about rule-following are famously also at the core of Kripke’s (1982) reading of Wittgenstein (1958). While the two sets of worries are not unrelated, Kripke’s primary focus is on meaning or content, not reasoning as such.
counts as following a rule just in case it manifests a suitable disposition of yours. The trouble arises in trying to be more precise about what such a disposition might be.

In the wake of Kripke’s reading of Wittgenstein’s remarks on rule-following (Kripke 1982), it is now generally agreed that a simple disposition to do \( x \) in circumstances \( C \) is not enough for you to count as following a rule that sanctions doing \( x \) in \( C \). To return to an earlier example, we are disposed to form beliefs about the emotional states of others on the basis of subtle facial and behavioral cues. Such beliefs, however, are generally fixed by automatic sub-personal processes, and do not count as instances of reasoning in the sense we are interested in here. We need to add something more, to distinguish between merely acting \textit{in accordance with} a rule and \textit{following} a rule.

The best developed version of the dispositional view I am aware of is Broome’s (2013; 2014). Broome suggests adding to the dispositional account a kind of \textit{seeming} to follow a rule, according to Broome, is to manifest a disposition to act \textit{in accordance with} the rule while this way of acting \textit{seems right} to you (Broome 2013, 237-8; 2014, 21). Seeming right in Broome’s sense is a higher-order attitude with relational content: an act seems right to you in light of a particular rule (Broome 2014, 22-3).

But why would adding this further state of seeming right help? Simply being \textit{accompanied} by a state of seeming right is not sufficient to turn a first-order disposition into an instance of reasoning. Consider the following possibility. Start with a “mere” disposition to act \textit{in accordance with} a rule: say, the disposition to believe you hear trumpets upon believing that it is raining. This, of course, is a rather bizarre disposition to have; but suppose it has been instilled in you through hypnosis. As Broome would accept, manifesting \textit{this} disposition is not sufficient for rule-following, and hence for reasoning. After all, even if you are disposed to believe that you hear trumpets upon believing that it is raining it will still (presumably) not seem right to you to do so. But now suppose that the hypnotist can from instilling a \textit{further} and \textit{distinct} disposition in you, this time with the effect that believing that you hear trumpets upon believing that it is
raining does seem right to you. Importantly, this disposition plays no role in the operation of your earlier disposition to believe you hear trumpets upon believing that it is raining: it is simply an independent feature of your psychology. It seems clear that merely adding this further disposition to your repertoire would do nothing to turn a transition that failed to count as reasoning before to an instance of reasoning now. (This point is structurally parallel to one that Broome 2013, 228 himself registers, to the effect that merely adding a higher-order linking belief to a process that is not reasoning will not turn it into reasoning. But if higher-order beliefs cannot achieve this feat it is hard to see how seemings could.)

Thus seeming right cannot merely accompany the first-order disposition to do $x$ in $C$. It must have a further role to play. Moreover, in light of our argument from the beginning of this section, we must make sure that whatever role it plays is not cashed out in inferential terms: Broome cannot, for instance, require that you infer from its seeming right to you to do $x$ in $C$ in light of $r$ that you should do $x$ in $C$. So how should we understand the role of seeming right?

Broome’s response, if I understand it correctly, is that the role of seeming right is essentially negative. What is essential to reasoning, according to Broome, is that it is open to correction: it essentially involves a higher-order disposition to change your mind under certain conditions. In order for your doing $x$ in $C$ to count as an instance of reasoning, in other words, it must be the case that there are conditions under which you would change your mind about it. Those conditions are conditions in which it no longer seems right to you to do $x$ in $C$ in light of some rule $r$ (Broome 2014, 22). This means that, according to Broome, following a rule does not require inferring what to do from the rule itself and a representation of your current circumstances. But it is not clear that this is enough to avoid the problems altogether.

Note that, unless much more is said about the conditions under which you are disposed to change your mind, “openness to correction” places no meaningful constraint upon reasoning: presumably just about any psychological disposition is such that it can be extinguished under some

---

8 I thank a referee for pointing out mistakes in my earlier reading of Broome’s account.
circumstances. The danger for Broome is that he might be unable to give a non-circular account of the conditions under which you should be disposed to change your mind if your doing \( x \) in \( C \) is to count as reasoning.

Broome does not attempt to give a full account of those conditions. He does, however, emphasize that you must be disposed to change your mind in response to the results of “checking” (Broome 2014 22-23). The problem is that it is hard to see how this could be cashed out otherwise than in terms of reasoning. Broome gives the example of checking my answer to a multiplication problem by comparing it to the multiplication tables. But if my earlier answer to the multiplication problem stops seeming right to me upon comparing it to the multiplication tables, isn’t that because I reason that it cannot be correct, since it differs from the answer in the tables? A brute disposition to change my mind upon having the relevant visual sensations, for example, would surely not count as checking. Thus Broome’s appeal to checking would seem to be just a covert appeal to reasoning.⁹ In other words, Broome does not succeed in giving a non-circular dispositional account of reasoning; all that his appeal to openness to correction achieves is push the problem one step back.

I have used Broome’s account to illustrate the difficulties, but the problem goes deeper than the specifics of any one account. The real difficulty is simply that the proposed rules of reasoning are the wrong type of thing to play the role required of them in reasoning. Although advocates of the rule-following model are quick to point out that the rules of reasoning are not the same thing as the rules of a formal system, it is hard to see how rules of reasoning could avoid being formal, in a roughly analogous sense. Rules of reasoning are, by necessity, rules that concern relations among representations — your beliefs or their contents — rather than the subject-matter you are reasoning about. So it is hard to see how you can genuinely follow — or

---

⁹ Broome also suggests that sometimes checking can simply consist in repeating the same process. But, surely, repeating the same process and getting a different result should not just brutally cause me to change my assessment of my earlier result. The divergence may well give me reason to reconsider that assessment, but that would be a case of further reasoning, not just a brute disposition.
take instruction from — those rules unless, in some sense, you are thinking about your own beliefs or their contents (even Broome, as keenly aware as anyone of the trouble with higher-order theories, introduces higher-order seemings to capture this idea). But, intuitively, when you reason your attention is on the world, not on your own beliefs or their contents. Thus there is always going to be a gap between what you need to be thinking about in order to genuinely follow a rule and the subject-matter of your reasoning.

Rule-following theories, therefore, face a dilemma. According to the first horn of the dilemma, what goes on in the higher-order side of the gap (i.e., your thoughts about rules and the responses they sanction) does not affect what goes on in the first-order side (i.e., your thoughts about the subject-matter of your reasoning). But then how is what goes on in the first-order side still a genuine case of rule-following? So we are led to the other horn of the dilemma, according to which what goes on in the higher-order side of the gap does affect what goes on in the first-order side. The trouble with this, of course, is that it is hard to see how it would work, except through further reasoning (as in our example from the beginning of this section, and even on Broome’s more sophisticated account). But the need to appeal to reasoning at this point shows that rule-following cannot be what we fundamentally do when we reason.

I cannot prove that no rule-following theory can successfully navigate these difficulties. But I think we have seen enough to motivate searching for an alternative approach, one which

---

10 There is a tradition, stemming from a reading of Wittgenstein’s (1958) remarks on rule-following, that seeks to avoid this dilemma by introducing a way of following a rule that does not rest on interpretation, i.e., a way of following a rule that is not mediated by the application of any further rules (McDowell 1984; Miller 2015). On Miller’s (2015) view, for example, reasoning in accordance with the rule (MP) may require no more than an intentional state with the rule as (part of) its content, and which combines with your beliefs that if Socrates is human then Socrates is mortal and that Socrates is human to cause you, in the right way, to believe that Socrates is mortal. Now, the challenge for any such view will be to spell out what “the right way” of causing beliefs is. Proponents of such views, following Wittgenstein himself, appeal to things like “practice”, “custom”, “training” and so on. Let us say that such a transition counts as reasoning just in case it is the manifestation of a certain kind of (perhaps socially scaffolded) cognitive skill
departs from the rule-following model. As we will see, such an approach makes it possible to smoothly accommodate the Taking Condition, and hence to make progress towards explaining the sense in which reasoning is an exercise of agency on our part.

4. A Semantic Approach to Reasoning

I take the premisses, conclusions and intermediate steps of reasoning to be contentful statements, not empty strings of symbols. Intuitively, understanding a statement involves knowing how it represents things as being, or what things have to be like for it to be true. It seems natural to analyze such knowledge in terms of possibilities, or ways for things to be. To understand a statement is to know which of the ways for things to be are such as to make it true. Coming to believe or accepting a statement involves ruling out possibilities in which the statement is not true. The approach I develop below is based on the familiar idea that the epistemic aim of reasoning is to reduce uncertainty about the world, via the elimination of alternative ways the world might be (see, e.g., Robert Stalnaker 1987).

Notice that since our topic here is human understanding with its familiar limitations, not all of the relevant “possibilities” or ways for things to be are possible worlds: ways for things to be by the reasoner. As it happens, I have much sympathy with this sort of view; I just wonder whether the appeal to a rich enough concept of cognitive skill does not render appeal to rules superfluous. Consider the inference from “roses are red” to “roses are colored”. How should we explain a normal English speaking subject’s capacity to reason in this way? As will become clearer below, my view is this: her understanding of these two statements enables her to see that there is no way for the former to be true while the latter is not. Her reasoning is explained by cognitive skills — specifically, skills of semantic evaluation. It is not clear that there is any need to appeal to rules of inference at all. Of course, some of those who appeal to rule-following in their account of reasoning also hope to explain the cognitive skills involved in understanding in terms of rule-following (e.g., Brandom 1998; Boghossian 2003). For them, the appeal to rules is not redundant. But if we do not share this ambition it is not clear we should follow them on this. I thank an anonymous referee for pressing me on this point.

11 This is not intended as a reductive account of belief in terms of ruling out possibilities. The point, rather, is to draw upon our intuitive grasp of belief to introduce some quasi-technical vocabulary that will be useful in what follows.
in our sense need not be complete or even logically closed. This allows us to accommodate the fact that you can understand or believe a statement without grasping all of its logical consequences (although there remains an important sense in which you are committed to all the logical consequences of what you believe). You can believe \( p \) and “if \( p \) then \( q \)” without thereby also believing \( q \), since ways for things to be in which \( p \) and “if \( p \) then \( q \)” are true but \( q \) is not are not automatically ruled out. So only a subset of these ways for things to be are real possibilities — the classical possible worlds, say.\(^{12}\)

Knowing what things have to be like in order for a statement to be true does not require being able to give an informative description of the relevant possibilities. It only requires being able to pick out the relevant possibilities upon considering them. Picking out a possibility as one that makes a statement true is not just a brute a-rational response, but rather an exercise of a cognitive skill, in the same sense that picking out Barack Obama from a crowd of people is an exercise of a cognitive skill (in the latter case, a skill of perceptual recognition). Just as my capacity to pick out Obama is plausibly constitutive of my knowing who Obama is, my capacity to pick out the possibilities in which a given statement is true upon considering them is constitutive of my knowing what things have to be like in order for it to be true — i.e., of understanding that statement.\(^{13}\)

\(^{12}\) Taking epistemic space to extend beyond the space of logical possibility is a relatively familiar way of treating deductive ignorance and error. See, for example, Ian Rumfitt (2008) and Mark Jago (2014). Any account of reasoning will need some way to represent deductive ignorance, and this approach seems natural if one wishes to avoid syntactic approaches.

\(^{13}\) This is consistent with propositional accounts of “knowing \( w.h. \)” (see Stanley 2011, chap. 2 for an illuminating overview). My knowing who Barack Obama is, for example, might consist in my knowing the relevant range of demonstrative propositions of the form “this is Barack Obama”, in the right perceptual contexts. A similar account is plausible for a subject’s knowing what things have to be like for a statement to be true: it consists in knowing, upon considering a relevant possibility that makes the statement true, that this possibility makes the statement true. One may wonder whether it is plausible that ordinary understanding involves knowledge of meta-linguistic propositions of this sort; but the charge of over-
What is it to consider a possibility? Suppose you are contemplating making a move in chess. You begin by noting how it will change the position of the pieces on the board and how these changes will affect the balance of threats among them. You will then consider different possible responses by your opponent, and then counter-responses on your part — perhaps going a few moves deep. These are all examples of considering possibilities. Good players will be thorough in their search of the space of possibilities, and efficient in deciding which possibilities are worth taking seriously. Worse players will be less so.

Different accounts of this type of activity are possible. What matters for present purposes is that the most fundamental way we have of considering possibilities must not itself be a matter of inference. This seems intuitively plausible: an experienced chess player, for example, can simply call to mind the possibilities afforded by a configuration of pieces on the board, without needing to derive them from the rules of chess. Developing such a chess-playing imagination is quite plausibly constitutive of becoming a skilled chess player. Such a capacity may take different forms. On one account that is prominent in the psychological literature, you consider possibilities by constructing mental models in working memory (Johnson-Laird 1983; Byrne 2007). Perhaps this involves offline sensory simulation, at least in some cases (Williamson 2008). While such capacities will of course involve extensive information processing, there is again no need to construe this as personal-level reasoning.

So how does all this help with reasoning and the Taking Condition? In the present framework, it seems natural to say that $p$ follows from a set of statements $R$ just in case there are no real possibilities in which all members of $R$ are true and $p$ is not (where, as mentioned earlier, the class of “real possibilities” is a subset of ways for things to be — the classical possible worlds, say). Taking it that $p$ follows from $R$ plausibly consists in ruling out all ways for things to be in

intellectualizing ordinary skills is one that propositionalists about “knowing why” have to address in any case.
which all members of $R$ are true but $p$ is not. (This is just another way of saying that it involves recognizing that committing to $R$ also commits you to $p$.)

Now suppose you already believe $R$, and hence all ways for things to be that are open to you are such that all members of $R$ are true. Still, even if $p$ really does follow from $R$, it is not guaranteed that all ways for things to be that are open to you are such that $p$ is true, since ways for things to be do not have to be logically closed. But notice that recognizing that $p$ follows from $R$ in this context just is coming to believe $p$, since it consists in ruling out all ways for things to be in which all members of $R$ are true and $p$ is not — which, in the present context, means ruling out all possibilities still open for you in which $p$ is not true. Thus this approach smoothly accommodates the Taking Condition. A subject infers $p$ from $R$ in virtue of recognizing that $R$ constrains the ways things might be so as to guarantee that $p$ is true.14 (Of course, subjects need not explicitly articulate thoughts of this complexity. They might, instead, express themselves entirely in the material mode: “$R$; so, $p$”.)

One might wonder whether I am not merely replacing the variety of rules recognized by the rule-following model with a single very general rule, based on the intuitive definition of validity for arguments. But it is important to see that this is not so. Rules of inference concern relations among representations. A rule of inference answers questions of the form, “what should I believe, given what I already believe?” The present account makes no reference to such rules. What the reasoner already believes simply constrains what possibilities are epistemically open to her. Her reasoning, then, consists in nothing other than her recognizing that her conclusion holds in all those possibilities. She does not need to consider relations among representations at all.

---

14 It bears noting that taking $p$ to follow from $R$ on the present approach is not just an input to a causal process that leads one to believe $p$ by reasoning from $R$. On the contrary, it is coming to believe reasoning from $R$. One might wonder whether this makes my account incompatible with familiar causal theories of action, thereby undermining its claim to capture the sense in which reasoning is active. I return to this point in Section 5.
The point is easiest to see in a case of an intuitively — thought not logically — valid inference. Suppose Alma is playing poker with Beth, and she reasons from “I have the ace of hearts in my hand” to “Beth does not hold four aces”. In believing the premiss of this inference Alma rules out all possibilities in which she does not have the ace of hearts in her hand. But suppose Alma has not until now considered whether Beth might hold four aces. Thus, while she is (in some sense) rationally committed to Beth’s not holding four aces, on our model this does not yet count as a belief of hers. When Alma considers the matter, of course, she immediately sees that there are no real possibilities in which Beth holds four aces while she holds the ace of hearts: this is simply an exercise of her capacity for recognizing possibilities that make statements she understands true or false. Thus any such ways for things to be are now ruled out for her. But this is exactly what it takes for Alma to recognize that it follows from her holding the ace of hearts that Beth does not hold four aces. But since Alma has already ruled out all ways for things to be in which she does not hold the ace of hearts, in ruling out ways for things to be in which she holds the ace of hearts and Beth holds four aces she comes to believe that Beth does not hold four aces. Her reasoning is done, without any application of a rule of inference.

A similar account, with some further assumptions, could work for non-deductive reasoning as well (non-deductive reasoning, being notoriously hard to codify, remains a problem for the rule-following model). Suppose that Raji sees Bob walk out of the examination room looking happy. She infers that Bob did well on his exam. Her inference is not deductive: even given her background folk-psychological knowledge Raji cannot rule out all possibilities in which Bob’s happy demeanor coexists with his having done poorly on his exam.\textsuperscript{15} Thus Raji cannot infer deductively that Bob did well on his exam. But suppose we are willing to grant that Raji knows that possibilities in which Bob’s happy demeanor co-exists with his having done poorly on the exam are, in some sense, far-fetched or abnormal, and that, absent any evidence to the

\textsuperscript{15} Bird (2005) denies this for cases of abductive inference. It would take us too far afield to consider Bird’s objections.
contrary, she is justified (perhaps by some sort of default entitlement, in the sense of Wright 2004b) in ignoring them. If all this is granted, then Raji is in a position to infer that Bob did well on his exam, by restricting her attention to non-far-fetched possibilities. Once again, notice that the Taking Condition is smoothly satisfied: Raji’s inferring that Bob did well on his exam just is her recognizing that all the possibilities that make her premiss true, and which additionally satisfy the assumption of normality, make her conclusion true.

So what is the point of formal rules of inference, on this account? Return to the modus ponens inference discussed earlier. Suppose our subject grasps and accepts the statements “if Socrates is human then he is mortal” and “Socrates is human”, and accordingly rules out all ways for things to be in which either of them is false. Given our assumptions this does not mean she automatically rules out all ways for things to be in which Socrates is not mortal: there is a further cognitive act she needs to perform in order to infer that Socrates is mortal. But, of course, even if she never performs this further act, it remains the case that she is rationally committed to Socrates’s being mortal. Moreover, her commitment depends only on how the original statements are put together, and not on anything specifically to do with Socrates or mortality. Recognizing the validity of the rule of modus ponens amounts to our subject’s recognizing this fact: it amounts to her recognizing that any statement of the form “if $P$ then $Q$” commits her to ruling out all ways for things to be in which $P$ is true but $Q$ is not, regardless of what is substituted for $P$ or $Q$ (at least for atomic $P$ and $Q$). This, I take it, is the point of formal rules of inference: they are not rules for reasoning, but rather for describing the logical structure of our commitments (a point also suggested by Gilbert Harman 1986).

5. **The Place of Reasoning in our Cognitive Lives**

In the last section I sketched a semantic approach to reasoning. But does this approach really provide us with a better understanding of the activity of reasoning? I would like to close this paper by explaining why I think so.
Let us begin with a question I shelved in Section 2: does my account capture the full sense in which reasoning is an exercise of agency on our part? Since I have no general theory of mental agency to hand, my answer to this can only be tentative. My account captures the sense in which reasoning is a way of making up your own mind: beliefs arrived at by reasoning are beliefs that reflect your take on your evidence. This is a form of control or agency that we have over our own doxastic states — evaluative control, in Pamela Hieronymi’s (2009) terms. The question is whether this is a robust enough notion of agency. So let us consider what more one might ask for.

My account does not capture any sense in which reasoning might be intentional or voluntary: for all that my account says, intention or volition might be relevant to reasoning only in the sense that we can intentionally choose to consider one question rather than another. But it is not at all clear that intention or volition ought to be given a bigger role than this in (theoretical) reasoning: believing on the basis of volition or intention is wishful thinking, not reasoning.

More interestingly, one might wonder whether my account is compatible with thinking of reasoning as a kind of process that we actively engage in. According to the “standard story” of action, physical actions consist of bodily movements that are controlled or guided by appropriate mental states (beliefs, intentions, plans etc.) that are distinct from them (Hornsby 2004; Smith 2010). According to the standard story, in other words, there is a kind of distance between the event an action consists in and the mental states in virtue of which that event counts as an action. But my account undermines this distance in the case of reasoning: if you take it that your premisses support p in the sense outlined above, that just is believing p by reasoning. Reasoning, then, is not a process that is controlled or guided by your assessment of the evidence; it just is your accepting a certain conclusion, by taking it to be supported by your evidence. (For similar points see Gibbons 2009; Boyle 2011; Valaris 2014.) As Alan White puts it: “inference is not the passage from A to B, but the taking of B as a result of reflection on A” (White 1971, 291; quoted by Rumfitt 2011).

16 I thank an anonymous referee for pressing me on this.
For present purposes the question is whether the claim that reasoning is not a process controlled or guided by your assessment of your evidence is in tension with the idea that reasoning is active. I do not think that it is. It is hardly clear that the distance posited by the standard story of action — the distance between the events that constitute our actions, and the mental states in virtue of which those events count as actions — really is an essential feature of our pre-theoretical conception of agency. But then the fact that my account eliminates this distance is no reason to doubt that it captures the sense in which reasoning is active.17

On a different front, an objector might point out that my approach liberally appeals to fairly sophisticated cognitive skills, such as skills for considering possibilities and evaluating propositions in them. But these cognitive capacities are, as the objector might reasonably claim, no less in need of an explanation than our capacity to reason itself.

More specifically, on my view reasoning is explained in terms of understanding and the cognitive skills that go along with it. But this way of approaching matters is incompatible with the “inferentialist” project in the theory of content and understanding, according to which understanding is to be analyzed in terms of inference — where the latter is typically conceived of in rule-following terms (Brandom 1998; Boghossian 2003). Although I cannot go into this in the depth that it deserves, I suspect inferentialists are simply getting things backwards here: for, how can inference be a rational phenomenon at all (rather than a brute reflex), unless you at the very

17 For a theory of mental action that attempts to reproduce something like the structure found in the standard story, see Proust (2009). Proust’s account is intended to apply not just to reasoning, but also to things like actively searching one’s memory. In this latter case it is not implausible that we really do have a process (searching one’s memory) which is controlled or guided by distinct mental states (e.g., your desire to remember someone’s name). By the same token, however, it seems doubtful that this external guidance shows that the process of remembering itself is an action: it is more natural to say that I actively bring it about that I remember than that I actively remember (Mele 2009 makes the same point). But then this structure cannot give us what we want in the case of reasoning: we want a sense in which reasoning itself is active, not simply a sense in which I can actively bring it about that I reason.
least understand what you are reasoning about? Of course much more needs to be said about this; but, at least prima facie, I do not think the inferentialist-inspired objection has much force.

More broadly, reasoning is a central example of the sort of control we have over our own cognitive lives. As such, it is a high-level cognitive skill. It is no surprise, therefore, to find that it works by drawing upon other, more specialized capacities, such as, for example, imaginative capacities or capacities for sensory simulation. Such capacities are, of course, highly complex in their own right: they draw on our understanding of common-sense physics and psychology, knowledge of chess, and more. They are certainly worthwhile topics for further study. But I think it is actually an advantage of my approach that it helps us see how our capacity to reason is constitutively connected with other cognitive capacities, including such high-level ones.

One important consequence of this fact is that the present approach can draw on a rich array of resources to explain patterns in reasoning performance. Consider, for example, the much-discussed fact that people find reasoning tasks easier when they are framed in familiar terms than when they are specified in unfamiliar, abstract, or nonsensical terms, even if the tasks are formally identical. From the point of view of rule-following theories, this fact must seem mysterious: shouldn’t reasoning simply abstract from content altogether? But if reasoning fundamentally involves the consideration of possibilities, then — in principle, at least — there need be no mystery here. The exact psychological mechanisms will need to be worked out empirically, of course, but in principle it is not surprising that people will find possibilities concerning familiar topics easier to think about than possibilities specified in unfamiliar terms.

---

18 Boghossian 2012, 17-18 recants his own earlier inferentialist leanings on similar grounds.
19 Much of the evidence for such “content effects” comes from research with the Wason selection task paradigm. For example, subjects perform much better with versions of the Wason selection task (Wason 1968) which are about familiar topics than with versions that involve either meaningless symbols or unfamiliar content (Wason and Shapiro 1971; Pollard 1981). Such content effects show that our reasoning capacities are, at the very least, not purely formal — they are not insulated from background knowledge and processes of semantic evaluation.
Now, its ability to explain such patterns in reasoning performance is the main selling point of the influential “mental models” theory of reasoning in psychology. Moreover, as I noted near the beginning of this paper (n. 3), this approach is often advertised as “semantic”, and is motivated by dissatisfaction with theories that postulate a special faculty of formal rules to explain reasoning. Thus it might be worth taking a moment to comment on the relation between this theory and the approach advocated here.

On one level, the claim that mental model theory is “semantic” is surely misleading. According to mental model theory, we reason by constructing in working memory models that interpret our premisses, and then checking whether those models provide counterexamples to our conclusion (Johnson-Laird 1983; 2001; 2008; Johnson-Laird and Byrne 1991). But mental models, no less than strings of symbols in a language of thought, are *syntactic objects*: they are intrinsically meaningless structures in the head that stand for worldly states of affairs. Moreover, the elaborate rules the theory postulates for handling the models are, clearly, formal: their application depends only on the properties of the models themselves, not on what the models are models of.\(^{20}\)

Now, as Johnson-Laird and Byrne (1993, 368) acknowledge, this is arguably inevitable, since mental model theory is a theory of how reasoning is computationally implemented and computational processes are not directly sensitive to semantics. (My proposal, by contrast, is a proposal about what *you* do when you reason.) But this leaves open the question of what exactly proponents of the theory mean when they call their approach “semantic”. From the present perspective we can see some point to this label: rather than postulating a self-contained faculty of rules, mental model theory analyzes our capacity to reason by showing how it draws upon other capacities, such as capacities for linguistic interpretation, imagination, and the consideration of

---

\(^{20}\) This is a familiar complaint, made by a number of participants in an Open Peer commentary on *Deduction* (Johnson-Laird and Byrne 1991) in *Behavioral and Brain Sciences* (Andrews 1993; Bundy 1993; Stenning and Oberlander 1993; ter Meulen 1993).
possibilities. In this respect the mental model theory is close in spirit to the approach recommended above.

In this paper I have tried to show that there are important philosophical reasons to abandon the view that reasoning, conceived as a personal-level activity, is fundamentally a matter of following formal rules. I suggested, instead, that we should think of reasoning in semantic terms. This, as I have tried to show, can result in a better understanding of the activity of reasoning and its place in our cognitive lives.21

References


21 I would like to thank three reviewers for Synthèse for their generous and constructive comments. Earlier versions of this paper were presented at the philosophy seminar at the University of Adelaide and the Minds Online 2015 conference. On the latter occasion I benefitted from written comments from Matt Boyle, Zoe Jenkin and Chris Tucker. I would like to thank the participants of both events for many insightful comments and suggestions.


