

What Reasoning Might Be

Markos Valaris

University of New South Wales

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 the rule-following model of reasoning, 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 paper suggests replacing the rule-following model with a 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.¹

What is the nature of this cognitive act?

¹ My attention here is restricted to reasoning with non-graded attitudes, such as knowledge and (full) belief. This choice is controversial, as some theorists argue that virtually all human reasoning involves graded doxastic states, or credences. I cannot enter this debate now, but it is worth noting that one reason for taking this attitude, namely that much of human reasoning is non-deductive in nature, has no force against the position to be argued for here: the approach I will sketch has no trouble accommodating non-deductive reasoning.

According to many, reasoning is fundamentally a matter of *following rules*.² 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 virtue of being guided by*, or *following*, a rule — a mental analog of the rule of modus ponens familiar from propositional logic.

The rule-following model is not often explicitly defended as such; most contemporary work on the nature of reasoning simply appears to take it for granted (see, e.g., Boghossian [2003; 2008; 2014], Broome [2013], Ichikawa and Jarvis [2013], Wedgwood [2002; 2006; 2007], Wright [20014]).³ 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 paradigmatic case of epistemic agency, or the kind of *control* that we have over 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.

Now, 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 (2014, 16-18). Boghossian, like other proponents of the rule-

² 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.

³ This is not to say that it has gone entirely unchallenged. Ian Rumfitt (2008; 2011), for one, proposes an alternative that is in many ways similar to my own.

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 the subject-matter of our reasoning (although some proponents of the rule-following model, such as John Broome [2006; 2013], would like to avoid “higher-order” conceptions of rule-following, as we shall see it is very doubtful that their attempts can succeed). 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.⁴

⁴The “mental models” theory (Johnson-Laird 1983; 2001; 2008; Johnson-Laird and Byrne 1991) in the psychology of reasoning is also often advertised as “semantic”. However, care is needed in interpreting this claim, as pointed out by a number of participants in an Open Peer commentary in *Behavioral and Brain Sciences* (Andrews 1993; Bundy 1993; Stenning and Oberlander 1993; ter Meulen 1993). This is because mental models themselves are, no less than sentences in a language of thought, *syntactic objects*, and the elaborate rules Byrne and Johnson-Laird describe for their manipulation are similarly purely formal as well. Discussing the mental model theory in detail goes beyond the scope of this paper. For present purposes, 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. Frege's 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 explain and motivate this condition.

Consider Frege's (1979, 3) characterization of inference, which Boghossian (2014, 4) also quotes approvingly:

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*. This is what Boghossian (2014, 5) calls the “taking condition” on inference, and what I will call “Frege's condition”. Is this 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, Frege's 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 Frege's 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 *agency* on our part; it is an exercise of the sort of control that we have over our cognitive lives. 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 very natural way to explain this difference is to say that inferring reflects *the subject's take* on what her evidence requires. By contrast, our system for judging other people's emotional states appears to be a lower-level system, whose workings are opaque to us. Marking this difference is the point of Frege's condition.⁵

Now, the question naturally arises as to what sort of state the "takings" required by Frege's condition might be. I would like to leave this question as open as possible, except for one important constraint. Consider the following case. Tom has some 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 is evidence that Bob 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. And yet Tom's belief that Bob will soon die is not, intuitively, justified, no matter how reliable a sign of impending death the spots might be. This is because, 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*

⁵ Perhaps it is possible to capture this difference without Frege's condition. Crispin Wright (2014, 33), for instance, acknowledges the need to "save the idea of inference as something that *we do*", while rejecting Frege's condition. He suggests, instead, that we should think of inference as a kind of intentional action. Wright, however, does not explain how he proposes to think of intentional action. This matters, because according to a familiar tradition in action theory, intentional actions are characterized by the agent's ability to answer the reason-seeking question "why?" (Anscombe 1957). If this is correct, then Wright's suggestion takes us back to Frege's condition.

them to support this conclusion.⁶ Thus the takings required by Frege's condition must exemplify states that can be assessed for epistemic justification.

The reason I emphasize this point is that it rules out views — such as those proposed by John Broome (2013) and Chris Tucker (2012) — that identify the relevant takings with *intellectual seemings*. Since seemings are not the sort of thing that can be either rational or irrational, such views would have trouble explaining what is wrong with Tom's inference above. After all, it surely *seems* to Tom that the spots on Bob's face are evidence that he will soon die, and this seeming is veridical.⁷

Now, the archetype of a state that can be assessed for epistemic justification and rationality is *belief*. This suggests that the takings required by Frege's condition might be beliefs. Since the nature of belief remains a controversial topic, this suggestion does not take us very far. Fortunately, more detail is not necessary for our purposes. I will occasionally speak as if the takings required by Frege's condition are beliefs but, so long as the point about epistemic assessment stays in view, nothing much hangs on this. (For example, views according to which the relevant takings are more akin to non-cognitive states of endorsing a norm — in something like the sense of [Gibbard 1986] — are not ruled out by this requirement, since such states are supposed to be open to rational assessment.)

Finally, one might wonder what exactly the *content* of the takings required by Frege's condition is. Once again, I want to leave this question as open as possible. Frege's condition

⁶ One might argue that Tom's theoretical beliefs should be construed as *extra premisses* in his reasoning. In that case the problem with Tom's reasoning would be that he is not justified in believing his premisses, rather than that he is not justified in taking his premisses to support his conclusion. I have no particular interest in defending this specific example, but the broader point should be resisted: there is an important distinction between *premisses* and *background knowledge* in reasoning, which the present objection would threaten to collapse.

⁷ Tucker (2012, 338) acknowledges the intuition that in cases like Tom's the subject's conclusion is unjustified, but he recommends simply setting it aside. Tucker gives no direct argument for this, other than that it is required by his own positive view. I suggest we do better keeping the intuition and rejecting the aspects of Tucker's view that conflict with it.

requires that, in inferring p from a set of premisses R , you must believe that R — in some sense — *supports* p . Different views might differ as to what exactly the relevant relation of support is, and even how exactly one must conceive of the *relata*. I will explain my own view in Section 4.

Let us now turn to the question whether the rule-following model is compatible with Frege's condition.

3. Frege's 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 Frege's condition holds, our subject must take it that (1) and (2) somehow provide justification for believing (3), and come to believe that Socrates is mortal partly because of this. Can rule-following theories explain how our subject meets this condition?

Let us begin by considering what the rule governing this bit of reasoning is. As most contemporary theorists recognize, this rule is not the familiar modus ponens rule of propositional logic: that is a rule that concerns 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 Frege's condition? It is hard to see how it could. 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 :

- 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 premises, 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. This is because 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.⁸

I have assumed so far that rules figure in our subject's thinking as the contents of *beliefs*. But this assumption may well be challenged: perhaps being committed to a rule of reasoning is a *sui generis* type of state, which is on the one hand assessable for rationality and on the other capable of directly motivating belief in the right circumstances. Thus, for example, upon coming to believe that if Socrates is human then he is mortal and that Socrates is human, a subject committed to the rule (MP) does not have to *reason to* the conclusion that she is rationally permitted to believe that Socrates is mortal — she need be in no doxastic state from whose content such a conclusion would follow. Rather, her commitment is manifested in the fact that

⁸ The history of this argument 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), Railton (2006), Broome (2013), Wedgwood (2006), and others. For some replies, see Leite (2008) and Valaris (2014).

she *is disposed to believe* in this way. Such a view, therefore, would seem to avoid having to explain reasoning in terms of further reasoning.

One problem with such views is that we seem to have no clear conception of what such non-doxastic states of commitment are. Suppose, first, that such states simply consist in dispositions to believe in the relevant ways. This is not satisfying, because a subject's disposition to infer in accordance with — say — (MP) would be in need of explanation, just as much as her actual modus ponens inferences. Thus the relevant states cannot be *mere* dispositions to believe, but rather states that *ground* and *explain* the subject's dispositions to believe in accordance with the relevant patterns — say, in accordance with the rule (MP).

Crucially, however, not just any sort of explanation would do: what we want is a kind of *rationalizing* explanation, an explanation that would show the relevant patterns of belief to be rational from our subject's point of view.⁹ Thus the states in question need to be states of a sort such that, in ascribing them to the subject, we partially specify her epistemic perspective on a certain subject-matter — namely, what would be rational for her to believe. Needless to say, paradigmatic states of this sort are beliefs — which, by hypothesis, the states in question are not.

But the real difficulty for the rule-following model in this area does not have to do with the precise nature of the states in question. 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. Consider again rule (MP) above. It is a rule about our subject's own beliefs. But our subject is not supposed to be reasoning about her own beliefs; she is supposed to be reasoning about Socrates and his mortality. It is therefore no surprise that we have trouble finding a place for (MP) in her reasoning.

⁹ Boghossian (2008, 498–9; 2014, 14–5) makes a similar point. Some authors (Peacocke 1995; Boghossian 2003; Wedgwood 2007) suggest that possession of certain concepts partially consists in dispositions to reason in certain ways. However, even if such dispositionalism is correct as an account of concept-possession (for arguments against see Williamson [2003; 2008]), it still does not give us the sort of explanation that we need here. Unless it is by way of explaining why belief patterns that conform to (MP) appear rational from our subject's point of view, it is hard to see how a subject's grasp of the conditional could explain her disposition to infer in accordance with (MP) in a satisfying way.

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, the above considerations show that rules of reasoning they propose tend to remain formal, in a very important sense. They are rules that concern relations among *representations* — your beliefs or their contents — rather than the subject-matter you are reasoning about. It is this formality that makes it hard to see how the rule-following model can accommodate Frege's condition. Intuitively, when you reason your attention is *on the world*, not on your beliefs or their contents (of course, your beliefs constitute your access to the world, and thus in reasoning you inevitably operate *with* your beliefs; but this is not the same thing as operating *on* your beliefs or their contents). Thus there is always a gap between what your reasoning is intuitively about and the manipulations of representations that the rule-following model predicts. Of course, a rather trivial bit of reasoning will typically bridge this gap: as we saw, a subject can reason from her grasp of the rule (MP) to the conclusion that she is rationally permitted to believe that Socrates is mortal (at least given certain assumptions about self-knowledge and the alignment of higher-order and first-order beliefs). But the need to appeal to reasoning at this point strongly suggests that following rules cannot be the most fundamental characterization of what we do when we reason.

Seen in this light, attempts to avoid the problem with the help of dispositions to believe or *sui generis* states of commitment simply come too late. They may avoid (by *fiat*) the circularity of having to appeal to reasoning in order to explain reasoning, but they do not address the more fundamental issue — namely, the fact the rules around which the rule-following model is based simply seem to have no rational bearing on what ordinary subjects usually reason about.

But do rules of reasoning on the rule-following model *have* to be formal in this sense? It is sometimes suggested that the problem with rules such as (MP) is that they are formulated as *higher-order* statements, i.e., as statements explicitly about beliefs. This is an important theme for John Broome (2013), for instance. Broome's own formulation of the relevant rule is this:

From p and (if p then q), to derive q . (2013, 234)

I think Broome is correct to want to avoid higher-order rules. Unfortunately, his formulation does not succeed. For, what is it “to derive q ”? From context, it looks like “to derive q ” just means to *come to believe q by reasoning*. But then, while Broome’s rule avoids *some* references to beliefs, it does not avoid *all* of them: it remains a higher-order rule in disguise.

There is a deeper lesson here: it is very hard to see what a *genuinely* first-order rule of reasoning would look like. A rule of reasoning is supposed to instruct you on what to believe, given what else you know or believe. How could a rule do this *without* talking about beliefs? Consider the following statement which, as a statement about propositions, might be construed as an attempt at capturing Broome’s (2013, 231-232) idea that reasoning is “operating on contents”:

The truth of any two propositions of the form p and ‘if p then q ’ is conclusive evidence for q .

Clearly, if this is going to count as a genuinely first-order statement, we need to understand the notion of evidence in play otherwise than in terms of “making it appropriate to believe” and the like. But then in what sense could this, or anything like it, be construed as a *rule of reasoning*? Such a “rule” would be impossible to follow, because it would not tell you what to *do*. Going first-order is not really an option for the rule-following model.

The same point holds in the case of rules that govern so-called material inferences, or inferences that are intuitively but not formally valid (and also, of course, in the case of inductive or abductive rules as well). Suppose, for example, that Alma reasons from “the roses are red” to “the roses are colored”. What might be the rule guiding Alma in her inference? The first-order statement:

For all x , if x is red then x is colored

is simply a universal generalization, not a rule of inference. It does not tell Alma what to do. What we need, rather, is a higher-order statement that would instruct our subject that she is permitted (or perhaps required) to believe that something is colored upon learning that it is red. Even rules of material inference turn out to be formal in the relevant sense.

All this suggests that we need a new approach in the theory of reasoning, one which departs from the rule-following model. The approach I will sketch never requires the reasoner to think about anything other than the subject matter of her reasoning. This means that rules of the sort we have been considering play no role in reasoning either. As we will see, this is what makes it possible for this approach to smoothly accommodate Frege's condition.

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 analyse 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 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 may well be a 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.¹⁰ So only a subset of these ways for things to be are *real possibilities* — the classical possible worlds, say.

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 arational 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 being able to pick out Obama is plausibly constitutive of my knowing who Obama is, being able to pick out 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.

Notice that this is consistent with propositional accounts of “knowing *wh-*” (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.

An important consequence of such an account is that my knowledge that *this* is Barack Obama on a particular occasion is not *inferred from* my general knowledge of who Barack Obama is: it is, rather, a constitutive part of that knowledge. (Of course, perceptual recognition requires extensive information processing. As explained in Section 2, however, this does not mean that it involves *reasoning*.) The same holds of a subject’s capacity to pick out the possibilities that make a

¹⁰ Using incomplete or even impossible worlds to model deductive ignorance is a relatively familiar, though controversial, approach. 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 like a natural alternative to syntactical ones.

statement she understands true: it is not *inferred from* her knowledge of what things have to be like in order for the statement to be true, but is rather constitutive of it.

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 Frege's 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. *Believing that p follows from R* plausibly consists in ruling out all ways for things to be in which all members of R are true but p is not. 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. In this context, coming to believe that p follows from R 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 Frege’s 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. (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 here whether I am not merely replacing the various formal rules of inference recognized by the rule-following model with a very general rule, based on the intuitive definition of validity. But this is not so. The point is easiest to see in a case of an intuitively — thought not logically — valid inference, say from “the roses are red” to “the roses are colored”. In believing the premiss of this inference, Alma rules out all worlds in which roses are not red. But suppose Alma has not so far as much as considered the question whether roses are colored. Thus, while she is of course (in some sense) *committed to* their being colored, 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 roses that fail to be colored while being red: this is simply an exercise of her non-inferential capacity to recognize possibilities that make statements she understands true or false. Thus any such ways for things to be are now ruled out for her. As discussed in the last paragraph, this is exactly what it takes for Alma to *recognize that* it follows from the roses’ being red that they are colored. But since Alma has already ruled out all ways for things to be in which roses are not red, in ruling out ways for things to be in which roses are red but not colored she thereby comes to believe that roses are colored. 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. 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 contrary, she is justified (perhaps by some sort of default entitlement, in the sense of Wright [2004]) 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 Frege's 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, however, 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 rule out ways for things to be in which both premisses hold but Socrates is not mortal. But, of course, it is an important fact about this particular example (and others like it) that this further cognitive act ultimately depends only on how the original statements were put together, and not on anything specifically to do with Socrates or mortality. This reflects a structural feature of the space of possibilities which a subject can come to recognize. In particular, a subject fluent with the conditional should, in principle, be able to recognize that a 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. This is a cognitive achievement for our subject, since on our model such possibilities are not automatically ruled out in virtue of having beliefs of that form. In this way our subject gains insight into the *logical structure* of the space of possibilities, of a kind which was not available to her before. This, I take it, is the point of the rule of modus ponens, and more broadly of formal rules of inference: they are not rules

for reasoning, but for describing the structure of our commitments (a point also argued 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 help us understand what the activity of reasoning fundamentally consists in? I can imagine a potential objector pointing out that my approach liberally appeals to fairly sophisticated cognitive skills, such as skills for considering possibilities and evaluating propositions in them. These cognitive capacities are, as the objector might reasonably claim, no less in need of an explanation than our capacity to reason itself. By way of a conclusion to this paper, I would like to say something about where, in my view, our capacity to reason should be situated in the larger picture of our cognitive lives.

Reasoning is a paradigmatic case of cognitive agency — a central example of the sort of control we have over our own cognitive lives. As such, it is a high-level cognitive skill. For this reason, it is no surprise to find that it works by drawing upon other cognitive capacities, such as imaginative capacities or capacities for sensory simulation. Such capacities are, of course, highly complex in their own right: they draw on our grasp of folk physics, folk 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 striking patterns in reasoning performance. Consider, for example, the much-discussed fact that people find reasoning tasks easier when they are specified in familiar terms than when they are specified in abstract or nonsensical terms, even if the tasks are

formally identical.¹¹ 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.¹² 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.

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.

References

- Andrews, Avery. 1993. "Mental Models and Tableau Logic." *Behavioral and Brain Sciences* 16 (02): 334.
- Anscombe, G. E. M. 1957. *Intention*. Cambridge, MA: Harvard University Press.
- Boghossian, Paul. 2003. "Blind Reasoning." *Aristotelian Society Supplementary Volumes* 177: 225–48.
- . 2008. "Epistemic Rules." *Journal of Philosophy* 105 (9): 472–500.

¹¹ 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 processes of semantic evaluation. Notice that the Bayesian program in the psychology of reasoning (Oaksford and Chater 2007), while rather different from the proposal sketched above in that it is set in a probabilistic framework, is also explicitly inimical to the *formality* of the rule-following approach: background knowledge and semantic understanding are obviously essential to it.

¹² Boghossian (2014, 12) illustrates the generality of reasoning through an example of modus ponens with propositions drawn from general relativity. Boghossian's point is that a typical subject will be able to recognize the validity of the argument even if she has no clue what the premisses or the conclusion mean. What Boghossian overlooks, however, is that subjects are actually likely to find this example *much harder* than, e.g., his own earlier one involving rain and wet streets, even though the two are formally identical.

- . 2014. “What Is Inference?” *Philosophical Studies* 169 (1): 1–18.
- Brewer, Bill. 1995. “Mental Causation II: Compulsion by Reason.” *Aristotelian Society Supplementary Volumes* 69: 237–53.
- Broome, John. 2006. “Reasoning with Preferences?” In *Preferences and Well Being*, edited by Serena Olsaretti, 183–238. Cambridge: Cambridge University Press.
- . 2013. *Rationality through Reasoning*. Chichester: Wiley Blackwell.
- Bundy, Alan. 1993. “‘Semantic Procedure’ Is an Oxymoron.” *Behavioral and Brain Sciences* 16 (02): 339–40.
- Byrne, Ruth. 2007. *The Rational Imagination: How People Create Alternatives to Reality*. London: A Bradford Book.
- Carroll, Lewis. 1895. “What the Tortoise Said to Achilles.” *Mind* 4: 278.
- Frege, Gottlob. 1979. “Logic.” In *Posthumous Writings*, edited by Gottfried Gabriel, Hans Hermes, and Peter Long, 1–9. Oxford: Blackwell.
- Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Lanham, MD: Rowman and Littlefield.
- Gibbard, Allan. 1986. “An Expressivistic Theory of Normative Discourse.” *Ethics* 96 (3): 472–85.
- Harman, Gilbert. 1986. *Change in View*. MIT Press.
- Ichikawa, Jonathan Jenkins, and Benjamin W. Jarvis. 2013. *The Rules of Thought*. OUP Oxford.
- Jago, Mark. 2014. *The Impossible: An Essay on Hyperintensionality*. Oxford: Oxford University Press.
- Johnson-Laird, Philip. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, MA: Harvard University Press.
- . 2001. “Mental Models and Deduction.” *Trends in Cognitive Sciences* 5 (10): 434–42.
- . 2008. *How We Reason*. Oxford: Oxford University Press.
- Johnson-Laird, Philip, and Ruth Byrne. 1991. *Deduction*. Hove, UK: Psychology Press.
- Johnston, Mark. 1988. “Self-Deception and the Nature of Mind.” In *Perspectives on Self-Deception*, edited by Brian McLaughlin and Amelie Rorty. Berkeley and Los Angeles: University of California Press.
- Leite, Adam. 2008. “Believing One’s Reasons Are Good.” *Synthese* 161 (3): 419–41.
- Oaksford, Mike, and Nick Chater. 2007. *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford: Oxford University Press.
- Peacocke, Christopher. 1995. *A Study of Concepts*. Cambridge, Mass.: A Bradford Book.
- Pollard, P. 1981. “The Effect of Thematic Content on the ‘Wason Selection Task.’” *Current Psychological Research* 1 (1): 21–29.
- Railton, Peter. 2006. “How to Engage Reason: The Problem of Regress.” In *Engaging Reason: Themes from the Moral Philosophy of Joseph Raz*, edited by Ray Wallace, Philip Pettit, Stephen Scheffler, and Michael Smith. Oxford: Oxford University Press.
- Rumfitt, Ian. 2008. “Knowledge by Deduction.” *Grazer Philosophische Studien* 77 (1): 61–84.
- . 2011. “Inference, Deduction, Logic.” In *Knowing How: Essays on Knowledge, Mind, and Action*, edited by John Bengson and Marc A. Moffett, 334. Oxford University Press, Usa.
- Stalnaker, Robert C. 1987. *Inquiry*. Cambridge, Mass.: A Bradford Book.
- Stanley, Jason. 2011. *Know How*. Oxford University Press.
- Stenning, Keith, and Jon Oberlander. 1993. “Nonsentential Representations and Nonformality.” *Behavioral and Brain Sciences* 16 (02): 365–66.
- Ter Meulen, Alice. 1993. “Situation Theory and Mental Models.” *Behavioral and Brain Sciences* 16 (02): 358–59.
- Tucker, Chris. 2012. “Movin’ on up: Higher-Level Requirements and Inferential Justification.” *Philosophical Studies* 157 (3): 323–40.
- Valaris, Markos. 2014. “Reasoning and Regress.” *Mind* 123 (489): 101–27.
- Van Cleve, James. 1984. “Reliability, Justification, and the Problem of Induction.” *Midwest Studies In Philosophy* 9 (1): 555–67.
- Wason, P. C. 1968. “Reasoning about a Rule.” *Quarterly Journal of Experimental Psychology* 20 (3): 273–81.

- Wason, P. C., and Diana Shapiro. 1971. "Natural and Contrived Experience in a Reasoning Problem." *Quarterly Journal of Experimental Psychology* 23 (1): 63–71.
- Wedgwood, Ralph. 2002. "Internalism Explained." *Philosophy and Phenomenological Research* 65 (2): 349–69.
- . 2006. "The Normative Force of Reasoning." *Noûs* 40 (4): 660–86.
- . 2007. *The Nature of Normativity*. Oxford University Press.
- Williamson, Timothy. 2003. "Blind Reasoning." *Aristotelian Society Supplementary Volume* 77 (1): 249–93.
- . 2008. *The Philosophy of Philosophy*. 1 edition. Malden, MA: Wiley-Blackwell.
- Winters, Barbara. 1983. "Inferring." *Philosophical Studies* 44: 201–20.
- Wright, Crispin. 2004. "On Epistemic Entitlement (I): Warrant for Nothing (and Foundations for Free)?" *Aristotelian Society Supplementary Volumes* 79: 168–211.
- . 2014. "Comment on Paul Boghossian, 'What Is Inference.'" *Philosophical Studies* 169 (1): 27–37.