Who is a Reasoner?*

[Penultimate draft. Final version forthcoming in *Inquiry*]

Abstract. This paper aims to make progress in understanding the nature of reasoning. Its primary goal is to spell out and defend a novel account of what reasoning might be, in terms of how reasoning contributes to settling (practical and theoretical) inquiries. Prior to spelling out this constructive proposal, however, the paper problematizes a very common picture of reasoning in an attempt to demonstrate the need for an alternative approach. The overarching argument of the paper is comprised of three stages. The first attacks the predominant conception of reasoning for its specious restriction to *fully conscious and explicit* episodes. The second stage offers a replacement for this faulty conception, according to which a reasoning agent is one who may be represented *as if* she were undergoing a fully conscious and explicit process. Finally, the last stage proposes one way to fill in the above schema, articulating and defending conditions for representing agents as reasoning fully consciously and explicitly. Each of the above stages is independent of the others: One may accept that reasoning need not be conscious and explicit but reject the schema proposed as a replacement; and again, the schema may be accepted but not the particular proposal for filling it in.

1. INTRODUCTION

This paper aims to make progress in understanding the nature of reasoning. It has two complimentary aims. One is to sketch and defend a constructive proposal: A novel account of what reasoning might be. It is not by accident or stylistic choice that the titular question is couched in terms of the identity of the reasoning *agent* rather than the nature of reasoning itself; a central contention of the paper is that the former question is prior to the latter. To understand what reasoning is, we need first to get clear on who can properly be regarded as reasoning agents. But

* Comments and suggestions by Olle Blomberg, David Enoch, Dan Friedman, Dave Jenkins, David Kovacs, and two anonymous referees for this journal greatly improved this manuscript. I am very grateful to all of them. Many thanks also to the constructive discussion of material from this paper by audiences at the 'Ration Action and Moral Agency' workshop at the University of Pavia, and the 'Reasoning and Agency' conference Tel Aviv University. Research for this paper was supported by the Israel Science Foundation (grant no. 381/23).

an equally important aim of this paper is the destructive prior step of attacking what, contrary to many philosophers' view, reasoning is *not*.

As a first step towards appreciating the flaws of the predominant position on the nature of reasoning and the merits of the alternative advanced in subsequent sections, it would be useful to have before us a rough initial characterization, free as much as possible from substantive preconceptions, of the target phenomenon. In broad and preliminary terms, then, reasoning is understood here as a pervasive psychological phenomenon ranging over different domains of thought and action. One central internal classification is between theoretical and practical reasoning. As an example of each – the first one theoretical, the second practical –, consider:

(1) If the government's new measures aren't effective, house prices will continue to soar. The government's new measures aren't effective. So, house prices will continue to soar.

(2) I shall make pesto sauce. In order to make pesto sauce, I must get basil leaves. So, I shall get basil leaves.

The following schematic formulations highlight the structure of each case:

(1) belief <*p*>
belief <if *p* then *q*>
So, belief <*q*>

and

(2) intention <I shall *V*>
belief <I shall *V* only if I shall *F*>
So, intention <I shall *F*>

As these formulations make clear, reasoning typically follows distinctive patterns, such as modus ponens and means-ends or instrumental reasoning, among various others. Here and elsewhere, reasoning is standardly understood as a patterned or structured transition between

mental attitudes, comprised of premises or *premise-attitudes* followed by forming a *conclusion-attitude*[1] inferred from the premises. Reasoning can therefore helpfully be thought of as a kind of mental process,[2] and the task of explaining the nature of reasoning becomes that of characterizing this process. The characterization developed below strives to be broad enough to capture both theoretical and practical varieties, while acknowledging the existence of important differences between them. Some such differences will crop up in the course of the discussion, though for the most part they will be glossed over. The question of what sets practical reasoning apart from theoretical reasoning will not occupy us in what follows so much as what the two have in common. As will become clear, the account is meant to capture both good (correct, rational) and bad reasoning.

The next section (§2) sets out the paper's main criticism of the dominant conception of reasoning as a *fully conscious and explicit process*. While the common restriction to conscious and explicit episodes of reasoning is itself familiar, it is rarely scrutinized carefully as it is here and as it deserves to be. §3 then articulates and defends an alternative picture of reasoning. This involves, first, introducing a schema of reasoning that does not require that reasoners conduct their business consciously and explicitly but only that they be representable as if they were so doing (§§3.1-3.2). The following sect. (§3.3) offers a particular way of filling in this schema. But understanding reasoning in these terms is *optional* in that the reader need not accept this picture along with other stages of the argument as a package deal (though she is strongly encouraged to do so!) Indeed, each stage of the paper's argument is independent of the others: One may accept that reasoning

[1] Some philosophers contend that practical reasoning can at least sometime conclude in *acting*. The idea is famously advanced by Aristotle (see *De Motu Animalium* 701a11-22, and *Nicomachean Ethics* 1139a 21-22 & 1147a 26-31). More recent defenses are Tenenbaum (2007), Fernandez (2016), and Dancy (2018). I have sympathies with this minority view, but substantiating it goes beyond the scope of this paper.

[2] Here and throughout, 'process' is intended in a thin sense that does not take a stand on the precise metaphysics of processes, and in particular does not insist on distinguishing processes from events. For the event/process distinction and its possible significance, see for example Hornsby 2012, Steward 2013, Stout 2018, and Levy 2020.

need not be conscious and explicit but reject the schema proposed as a replacement; and again, the schema may be accepted but not the particular proposal for how to fill it in.

2. THE PREDOMINANT CONCEPTION OF REASONING AND ITS FLAWS

2.1. The fully conscious and explicit conception

We start with rough, slogan-form statements of some extant accounts illustrating the target conception of reasoning attacked in this section. John Broome (2013, chs. 12-13 especially), for example, suggests that reasoning is essentially a rule-governed process of attitude revision. Paul Boghossian is likewise sympathetic to the idea that reasoning is rule-governed. Having particularly *theoretical* reasoning in mind, he claims that for S to infer *q* from *p* "is for S to judge *q because* S *takes* (the accepted truth of) *p* to provide (contextual) support for (the acceptance of) *q*." (Boghossian 2019: 110, emphases in original; see also his 2014). Meanwhile, McHugh & Way (2018a) propose that reasoning is a functional process that constitutively aims at "getting things right", i.e. at getting correct, right, or as they put it "fitting" attitudes.[3] For example, on their view, theoretical reasoning may be said to constitutively aim at forming true (and therefore correct or fitting) beliefs, while practical reasoning may be said to constitutively aim at forming permissible intentions. Finally, Eric Marcus (2021) takes theoretical reasoning (inference) to be a self-conscious act in which the reasoner represents (hence, takes) her premises to support her conclusion.

While these accounts of reasoning differ significantly in detail, it is a fundamental general assumption they all share, along with various other writers, which sets them up for the criticism to follow. This is the idea that the process of reasoning is *fully conscious and explicit*.[4] To see what this

[3] Like Broome and Boghossian (and many others), McHugh & Way are also sympathetic to the thought that reasoning is rule-governed. However, as they explain, the rule-governed picture is a useful yet optional component of their view (2018: 182). Since rule-governess is not part of the following objections to the paradigm that the views surveyed subscribe to, the issue is set aside.

[4] Some writers, in attempting to account for reasoning, accept that it is explicit and consciously accessible yet do not see it as a protracted *process* at all. Thus for example, Valaris (2017) and Neta (2013) propose that inference is a matter of judging or accepting a certain conclusion, by taking it to be supported by

might mean, start with explicitness. The general approach that forms our target here seems to treat it as a pre-theoretical datum that reasoners must register the premise-attitudes and conclusion-attitude more or less as they appear in the formulations given in §1 above. On this understanding of explicitness, if for example one's reasoning includes some premise-belief with the content that there is a ball in the northwest corner of the room, and another-premise attitude with the content that there is a ball in the southeast corner of the room, then one's reasoning may be said to implicitly but not explicitly represent that there are *two* balls in the room. Explicitness in the intended sense also involves having exactly the attitudes in question (belief, say) and not some surrogate attitude that falls short of it (e.g. acceptance).

Further, on the predominant conception, the steps in one's reasoning are understood to be not only explicit but also consciously accessible. The reasoning agent is thought to be *consciously aware* (or at least, can become aware) of the premise-attitudes, the conclusion-attitude, and in some way or another also of the *support* the premises provide (as she sees things, anyway) for the conclusion.[5]

2.2. Is reasoning everywhere conscious and explicit?

Proponents of the predominant conception exclude cases that fall short of being fully conscious and explicit. But the restriction is highly problematic, since there are plenty of cases that do not meet this standard and yet are intuitively recognizable as reasoning, as will be illustrated now. This objection itself is familiar, and its target-proponents occasionally even respond to it in

one's evidence or premise-attitudes. These accounts see reasoning as having a somewhat different temporal profile from that of the process envisaged here (and by most other philosophers): instead of an unfolding process that progresses gradually through premise-attitudes before culminating in the conclusion-attitude, they see reasoning as a judgment with a particular sort of content.

[5] There is some controversy among those who subscribe to the conscious-and-explicit-conception over how exactly the support provided by the premises must register in one's reasoning. For more on this, see discussions of the 'Taking Condition' on reasoning (e.g. Boghossian 2014, Broome 2014, Wright 2014, Tucker 2012, McHugh & Way 2016, Valaris 2017). The issue will not be taken up here, primarily for reasons of space.

print. But it is an instructive exercise (one that has not, to the best of my knowledge, been attempted before) to engage more closely with the responses and reveal why they are in fact inadequate.

First, some examples. Consider a striking result from a famous study in social psychology. Gergely, Bekkering & Király (2002) showed that 14-month-olds would use their head rather than their hand to turn on a light panel, even though it was less convenient to do so, but *only if* they had grounds to believe that this is how you turn it on. Infants were shown a video of an actor turning on the panel with their head, in two different conditions: In condition (a) the actor's hands were seen to be occupied, while in condition (b) her hands were visibly free but she nonetheless still used her head. Gergely et. al. found that, when given the opportunity, only subjects in (b) imitated the actor. Now, it is very natural to see the infants here as reasoning along something like the following lines:

> That person turned on the light with their head. If she could, she would turn it on with her hand. So, one must use one's head to turn on the light.

The above reconstructs a chain of *theoretical* reasoning that could be attributed to the infants. Alternatively, they may be thought to engage in a piece of practical reasoning, concluding in an intention to turn on the light panel. Either way, the point to note is just how natural it is to portray the infants as engaging in reasoning, and how readily this helps to make sense of their behaviour. Attributions of reasoning in this and similar cases are also overwhelmingly commonplace among psychologists. Yet if there is genuine reasoning here, it falls short of the standard of full conscious accessibility and explicitness. At least some of the steps in the infants' reconstructed chain of reasoning are likely to be unconscious – and equally likely implicit, seeing as they will have lacked some of the requisite concepts.[6]

[6] Recall that implicitness in the intended sense can manifest in a roundabout and/or enthymematic representation. Thus for example, perhaps an infant who lacks the concept MUST might nonetheless

Next, consider an imaginary case from Susanna Siegel (2019: 18), which likewise seems to involve less than fully conscious and explicit reasoning, this time in adults:

> *Pepperoni.* Usually you eat three slices of pizza when it comes with pepperoni. But tonight, after eating one slice, you suddenly don't want any more. Struck by your own uncharacteristic aversion, you form the belief that the pizza is yucky. Though you don't know it, you're responding to the facts that (i) the pepperoni tastes very salty to you, (ii) it looks greasy, (iii) it reminds you of someone you don't like, who you recently learned loves pepperoni, and (iv) you have suddenly felt the force of moral arguments against eating meat. If the next bites of pepperoni were less salty, the greasy appearance turned out to be glare from the lights, you learned that your nemesis now avoids pepperoni, and the moral arguments didn't move you, the conclusion of your inference would weaken, and so would your aversion. You haven't classified what you see and taste as: too greasy, too salty, reminiscent of your nemesis, or the sad product of immoral practices. Nor are you consciously thinking right now about any of these things.

Once again, and as Siegel herself points out, *Pepperoni* seems intuitively to be a case of reasoning. There are features of the pizza (its greasy look and saltiness), features of yourself (being reminded of your nemesis), and moral features (the moral status of eating meat) that you are responding to in forming the conclusion-attitude (a point we shall come back to below). Moreover, your conclusion is evaluable for its degree of (ir)rationality or justification – arguably irrational/unjustified with respect to the first two considerations, (more) rational with respect to the third (Ibid.) Yet once again, the case would not count as reasoning according to the predominant conception, since it is hardly fully conscious and explicit.

Finally, consider a famous study of animal cognition (Weir, Chappell & Kacelnik 2002). A New Caledonian crow named Betty was presented with some food stored inside a small basket at the bottom of a vertical plastic tube. On previous trials, Betty was given a hooked wire, and

reason in the way suggested in the text by deploying instead the concepts CANNOT (use her hand) and CAN (use her head).

managed to use it to lift the basket and retrieve the morsel. But when she only had available a straight wire, and after trying and failing to lift the basket with it, Betty took the straight wire to a fracture in a nearby plastic tray, and bent it there to form a hook. She then came back and retrieved the food.

Here, opinions about our question are likely to diverge. Some find it perfectly plausible to see Betty as reasoning while others would deny this, opting instead for explanations of her behaviour in terms of associative conditioning. The issue is closely tied to the broader possibility of animal cognition: Are some non-human animals capable of thought at all, and if so, which? The topic is hugely complex and goes well beyond the scope of the present discussion. However, it seems unwise to close off the very possibility that some non-human animals can reason purely on the basis of a definition of reasoning as necessarily conscious and explicit (a definition which Betty's behaviour is unlikely to meet). The question should rather be decided as part of a wider investigation into animal cognition, which examines the existence of other interlocking cognitive capacities, such as concept acquisition and deployment, perception, memory, communication, etc. etc. Indeed, this more holistic approach is increasingly how scientists and philosophers are approaching the topic (Bekoff, Allen & Burghardt 2002; Hurley & Nudds 2006; Lurz 2009; Tye 2016), so going against it would, again, break ranks with standard scientific (and philosophical) practice.

2.3. Some responses

As noted already, advocates of the predominant conception of reasoning are aware of the general line of objection rehearsed above, and some respond to it albeit briefly. For example, Boghossian (2014: 2-3), making reference to the widely accepted distinction between 'System 1' and 'System 2' information processing, explains that

> [w]hen I say that I'm interested in inference, I mean that I am interested in reasoning that is System 1.5 and up. That is to say, I'm interested in reasoning that is person-level, conscious, and voluntary, not sub-personal, sub-conscious and automatic.

In situating the phenomenon he is interested in closer to System 2, Boghossian means to characterize it as person-level, conscious, and voluntary yet not necessarily effortful and demanding, as System 2 thinking is typically thought to be. (See also McHugh & Way 2018a: 168 for a similar restriction.) Circumscribing his target in this way yields a well-defined phenomenon with philosophically interesting features. Why should that be considered problematic?

In answering this question, we should not let the discussion deteriorate into a verbal dispute over the correct reference of 'reasoning'. Nor is the issue approached here as an exercise in conceptual analysis. But that does not mean that Boghossian and others are free to circumscribe the scope of reasoning by stipulation if they are after an adequate characterization. We are faced with a natural (psychological) kind[7] – however it should be referred to – so getting its scope right is subject to natural constraints. As we have seen above, a restriction to fully conscious and explicit episodes squares badly with these constraints. It rules out intuitively recognizable cases, and fails to align with standard scientific and philosophical practices. Adopting the restriction therefore risks overlooking significant common features and affinities between cases that do, and cases that do not, meet the restriction. The resulting approach is impoverished in explanatory power.

[7] Why suppose that reasoning is a natural-psychological kind? One important piece of evidence is the commonplace use of the concept REASONING in the psychological and cognitive sciences. Moreover, if the supposition is mistaken and to circumscribe reasoning is not as it were a way to carve psychological nature at one of its joints, then it strikes me that the project of explaining what reasoning *is* would lose much of the interest it clearly holds. After all, why should so many talented philosophers be animated by the prospect of understanding the nature of some artificial or gerrymandered phenomenon? Finally, the argument in the text does not strictly require us to assume that reasoning is in fact a natural kind. All we need to acknowledge is that the narrower conception of reasoning as everywhere conscious and explicit is much less theoretically fruitful, as explained in the text. My thanks to an anonymous reviewer for getting me to consider this issue.

It might be responded that even if 'reasoning', understood more broadly as it is here, does denote a natural kind, this does not force one to circumscribe the target phenomenon accordingly. For perhaps the narrower predominant conception *also* denotes a natural kind. After all, what's to rule out natural kinds whose members are subsets of other natural kinds? Moreover, proponents of the predominant conception need not of course suppose that partly unconscious and implicit reasoning is *wholly unconnected* to the 'System 1.5 and up' variety. They could recognize the affinities between the different kinds of case and attempt to explain them indirectly – that is, not by classifying the unconscious and implicit kind together with reasoning proper. One way to read the above critique is therefore as mounting a challenge for champions of the predominant conception, viz. the challenge of finding a place within their approach for unconscious and implicit episodes. Recognizing the reality of this variety of reasoning is of course a precondition for meeting the challenge – one which at least Broome, Boghossian, McHugh & Way, Marcus, and others fail to satisfy. But even those adherents of the dominant view who *would* be willing to broaden their target conception would arguably be at a disadvantage compared to the approach recommended here, as their account would have to be considerably more complex and hence less attractive. Where they would propose (at least) two natural kinds with various connections, similarities and differences obtaining between them, the present account offers one general characterization encompassing both.[8] In any event, the project pursued here would boast a theoretical achievement of interest if it manages to subsume the narrower natural kind under a broader one – which the predominant conception is silent about.

A different (though related) response to the above criticisms highlights the reasons for restricting reasoning to fully conscious and explicit episodes. Two such reasons are the respective claims that (a) reasoning is active or agential (e.g. Broome 2013, chs. 12-13); and (b) that we can

[8] As will become clear in §§3.1-3.2, the characterization of reasoning proposed here does not eliminate the difference between the implicit and explicit kinds. Indeed, it gives *priority* to the latter when it comes to explaining the nature of reasoning. However (as an anonymous reviewer helpfully pointed out), the priority in question is methodological and not metaphysically substantive.

be held responsible for our reasoning (e.g. Boghossian 2019). Neither feature is thought possible unless reasoning is fully conscious and explicit. The two features (a) and (b) are sometimes noted in conjunction: it is allegedly *because* reasoning is active that we can be held responsible for it (cf. Boghossian 2019: 111, and Valaris 2017: 2010).

However, a host of open questions stand in the way of accepting activeness and responsibility as supporting a restriction to conscious and explicit cases, at least absent further argument. Consider: why think that reasoning is active in the first place? It is not enough to point out that reasoning is something 'we do' (Broome [2013: 208]; Boghossian [2014: 5]; McHugh & Way [2016: 314]), since waiting and resting are also things we do. (In general, the applicability of 'something one does' is a better guide to that something's being *personal-level* – a status happily accepted here as accruing to reasoning.) Brute intuition can hardly be regarded a dependable guide to agency either, as can be seen from the fierce controversy over the scope of mental agency (O'Brien & Soteriou 2009) – if there even *is* something worthy of that title (Strawson 2003, Metzinger 2013, Levy 2019). Doubts about the agential status of reasoning in particular can and indeed have been raised – for example, doubts stemming from the limited *control* we seem to have over our inferences (Jenkins 2021). It is of course open to friends of the predominant conception to defend some criterion for agency, a task they typically do not attempt, on which reasoning comes out as active. But even if that were accomplished, it is far from obvious that a defensible criterion would deliver the result that *only* conscious and explicit reasoning is active; after all, we presumably perform many habitual and other non-intentional actions (e.g. doodling or shifting position) 'on auto-pilot', without full conscious awareness. In any event, unless and until such a criterion is spelled out, it cannot simply be assumed that conscious and explicit, *and only* conscious and explicit, reasoning is active.[9]

[9] It should additionally be noted that, should the requisite criterion for agency be provided, any account of reasoning that rests on it would of course be hostage to its fortunes.

Turning briefly to the question of responsibility: Here, too, the matter is hugely complex and no firm conclusions can be reached in this space. However, precisely one upshot of this complexity is that one cannot simply assume that responsibility attaches only to conscious and explicit attitudes and events/processes. For one thing, so called 'attributionists' about moral responsibility (Scanlon 2008, Hieronymi 2008, and Smith 2005, among others), who hold, very roughly, that attributions of responsibility are grounded in the agent's evaluative judgments and attitudes, deny this suggestion. As Smith for example explains (2005: 252): "'Judgments' in [the intended] sense do not always arise from conscious choices or decisions, and they need not be consciously recognized by the person who holds them." Notice that the question of whether reasoning is *active*, as discussed above, should be neither here nor there when it comes to the aptness for attributions of responsibility. Such aptness need only require an understanding of reasoning, which it readily receives here and in most other places, as *person-level* in nature.

To summarize, the predominant characterization of reasoning as fully conscious and explicit is highly problematic, as it excludes intuitively recognizable cases that fall short of this standard and is discontinuous with scientific and philosophical practice, rendering its explanatory reach impoverished. Moreover, the reasons adduced for the restriction by its proponents are uncompelling: it is unclear that, nor why, reasoning should be considered active at all, and unclear that, nor why, *only* conscious reasoning should be so considered. That reasoning is apt for attributions of responsibility is again questionable grounds for thinking that it must be conscious and explicit, and indeed some central views of moral responsibility square badly with this idea.

The above makes no claim to have conclusively refuted the standard conception of reasoning as a necessarily fully conscious and explicit process. Hopefully however, absent counterarguments at least, enough doubt has been cast on it to merit exploring more promising alternatives.

3. WHAT REASONING MIGHT BE

3.1. Reasoning and representability

In rejecting the claim that reasoning necessarily follows, in conscious and explicit fashion, such patterns as laid out above (pp. 1-2), the alternative conception proposed here does not abandon these patterns; on the contrary, it makes essential reference to them. The core idea driving the proposal is that *a reasoning agent is one whose behaviour can be represented as if she were undergoing a fully conscious and explicit process of the sort envisaged by the predominant conception*. The proposal is designed to capture the idea that reasoning *may* be, but is not necessarily, conducted fully consciously and explicitly (someone who is in fact reasoning fully consciously could obviously be represented as if she were doing so). An important part of developing this proposal consists in the attempt below (§3.3) to identify *minimal* conditions under which one can be represented as if undergoing a fully conscious and explicit process of attitude formation/revision. These conditions are designed to capture also less than fully conscious and explicit cases, and are put forward as necessary and sufficient for someone to count as reasoning. But before those conditions can be stated and assessed, we need to gain a clearer view of what the relevant notion of representability involves. In doing so, we'd do well to take a cue from philosophy of science, where similar issues have been extensively investigated.

Vehicles of epistemic[10] representation facilitate formation of hypotheses about the target-systems they stand for. In developing this idea, one very general question that arises is what makes this practice possible. In virtue of what does some vehicle count as an epistemic representation of its target? Several different accounts are defended in the literature on scientific representation, which we need not pause to adjudicate between. Very briefly, *stipulative* or *denotational accounts* offer a deflationary answer, according to which all it takes for a vehicle to constitute an epistemic

[10] Epistemic representation is the genus of which scientific representation is a species. It denotes any representation done in the service of some inquiry, scientific or otherwise, and so excludes for example *aesthetic* representation.

representation of a certain target for a certain user is that the user stipulate that the vehicle denotes the target (Callender & Cohen 2006). *Inferential accounts* of epistemic representation build on stipulative accounts, adding the further condition whereby the vehicle must allow its user to make specific inferences about the target (Suárez 2004). And *interpretational accounts*, inspired by inferentialism, hold roughly that "a vehicle is an epistemic representation of a certain target (for a certain user) if and only if the user adopts an interpretation of the vehicle in terms of the target" (Contessa 2007: 57) – that is, iff the user sees constituents of the vehicle as standing respectively for parallel constituents of the target.

All three major conceptions of epistemic representation just listed are compatible with thinking of the fully conscious and explicit conception of reasoning as a vehicle for representing reasoners' behaviour, and hence either conception could be adopted for present purposes. Sentence-sequences such as (1) and (2) above (p. 1) can be seen as *denoting* patterned thought-processes attributable to reasoning agents, thereby facilitating *inferences* about those processes, which might also involve adopting some more specific *interpretation* thereof.

A closely related yet more delicate question than the one about conditions of epistemic representation as such concerns the conditions of good or *faithful* representation. (The two questions are not always disentangled.) Faithfulness in the intended sense is a gradable property: the better (more faithful) the representation, the more sound inferences it allows to draw about its target (though it need not be an *accurate* representation.) Here too, extant theoretical options – citing relevant similarities (Teller 2001; Giere 2004) or structural morphisms (da Costa & French 1990; Bueno, French & Ladyman 2002) between vehicle and target – are broadly compatible with the view of reasoning being developed. Importantly for us, according to all these accounts, what makes for a faithful representation will depend to a large extent on the particular *aims* the representation is serving. Consider an example: tractors in a certain village are equipped with a stronger reverse gear compared to their forward gear, making uphill climbs easier for the tractors

to make in reverse. Exploiting this fact, the village children are in the habit of representing the tractors as being afraid of going up steep slopes and so better at climbing them in reverse, when they could not see what lay ahead.[11] Is the children's representation faithful? That depends *inter alia* on what (if anything) they were trying to achieve. If the vehicle (no pun intended) was meant to predict how quickly and smoothly the target would travel uphill, the answer may well be that the representation was at least sufficiently faithful. But if the children were hoping to gain a deeper understanding of the inner workings of tractors, say, then probably much less so.

The lesson to draw is that to assess the adequacy of representing reasoners as undergoing a fully conscious and explicit mental process requires a clear idea of the aims to which this representation is being put. And in the present case, that aim is plausibly to make sense of agents' behaviour. Attributing sequences of reasoning to agents is of course a ubiquitous and highly efficient means for understanding why they act (believe, feel, etc.) as they do. It is here that we can begin to appreciate a throwaway remark made right at the start in connection with the title of this paper (p. 1). The point of focusing on what it takes to represent someone as reasoning comes from the priority this move accords to the *third*-person over the *first*-person perspective: If we wish to gain traction on the nature of reasoning, it is argued, we must first ask what it is to attribute reasoning to someone.

The reader might feel that this move is a bit like using a sledgehammer to swat a fly: If the central flaw of extant accounts of reasoning is their unduly narrow focus on conscious and explicit episodes, must the fix come from wielding in the heavy (and controversial) machinery involved in prioritizing the third-person perspective? Why not simply accept the reality of unconscious and implicit episodes while remaining firmly within the more conventional first-person outlook? In reply, the first point to note is that given the deliberately modular structure of the overarching argument of this paper, it is indeed possible to forgo the restriction to conscious and explicit

[11] Thanks to David Enoch for this example.

episodes while rejecting the further step of adopting the spectatorial perspective. However, if we stop there, we shall lose recourse to the explanation for *why* reasoners are typically represented in fully conscious and explicit terms. That explanation, already noted, is the straightforward one that explicit reasoning sequences are a common and perspicuous tool to make sense of agents' behaviour. We represent reasoners as reasoning consciously and explicitly, even when they are not, because that helps us see better (or show others) what they are up to and why. This seemingly innocuous observation already accords priority to the third-person outlook on reasoning developed in these pages.[12]

To see a bit more clearly how this approach manages to bear the explanatory payoffs touted, it may help to revisit briefly one of the examples from §2.2. Consider again the findings of Gergely and colleagues (2002). It was pointed out above that the infant subjects in this experiment are naturally seen as reasoning that one must use one's head to turn on the light in the video they were shown, because otherwise the protagonist would have used their hand. Understood as reasoning thus, it was suggested, helps us make better sense of the subjects' behavior. But how is this supposed to work exactly? When attributing to the infants a form of reasoning they did not actually preform (because they lack the full mental repertoire required to pull it off), are we not merely suggesting that they behave *as if* they were genuine reasoners? No. The claim is that the infants are in fact reasoning, though perhaps partly unconsciously and implicitly. They are reasoning because their thinking can be represented in fully conscious and explicit terms. And this makes available a powerful conceptual tool for understanding what they are up to – namely, the tool of representation in conscious and explicit terms.

A concrete proposal about the conditions for attributing conscious and explicit reasoning is made below (§3.3). But first, we pause to relate structurally similar explanatory strategies.

[12] I'm indebted to an anonymous referee for discussion here.

3.2. Quasi bedfellows

The general approach of seeking illumination by exploring conditions of epistemic representation is exploited by philosophers investigating other phenomena. One instructive analogy is with interpretationist approaches to mentality and in particular mental content. Consider Daniel Dennett's *Intentional Stance* [IS] (Dennett 1987, 2009), which famously is a strategy of interpreting the behavior of entities – persons, non-human animals, artefacts, etc. – by treating them as if they were minded, rational agents. To adopt IS towards some entity is to assume that that entity holds beliefs and desires which govern its behaviour. Moreover, the entity's 'beliefs' are assumed to be formed largely on the basis of the available evidence, and its 'desires' to be directed largely at achieving what is good for it. Finally, according to Dennett's *Intentional Systems Theory*, to be interpretable via IS *just is* to have the mental attitudes in question; there is no further question of whether the entity *really* or only *as-if* believes that *p* or desires to *V*. The program of 'radical interpretation' proposes a broadly similar tack for attributing mental attitudes to some agent as products of a bystander's interpretation of her behaviour, though somewhat less radically, it focuses primarily on *human* agents.[13] Fundamentally, what the present view shares with interpretationism is a plea, already remarked on, to reorient the lens through which certain mental items are standardly explored away from the first-person to the third-person perspective.

Classical interpretationism is rightly celebrated as an ambitious and fruitful contribution to philosophy of mind and cognitive science. But it is also open to various sources of puzzlement and protest. One persistent bone of contention is the sort of deflationary attitude towards mindhood that is at the heart of IST and may also be favoured by Davidson (see 1983: 315). Importantly, no such attitude is adopted here, and consequently the approach taken is much more conservative by comparison. It involves no commitment to anti-realism or instrumentalism about

[13] For classic defenses of interpretationism see Davidson 1973, and Lewis 1974; and more recently Pautz 2013, and Williams 2020.

the mental in general or about reasoning in particular. However, there may be other grounds to be suspicious of the interpertationist mode of explanation, which seem to apply also to its affiliate defended in these pages. To wit: focusing on the third-person observational perspective may appear to leave no room for the essentially first-personal character of one's mental goings-on.[14] Put in terms of the present phenomenon of interest, the complaint is that the agent is cast as a mere bystander to her own reasoning.

One way to reply would be to insist that interpretation is not restricted to the third person but is also a first-personal process. In reasoning, the agent goes through a process of self-interpretation.[15] This line of response may appeal to followers of Gilbert Ryle, but it is not recommended here.[16] A better way to see past this important objection proceeds by noting that the interpretationist framework works by narrowing down, as it were, the gap between third- and first-personal outlooks. While this is not the place to spell out a detailed account, it is a familiar contention of interpretationists that in order to represent some agent as having a specific mental state or as undergoing a specific mental process, the representor must deploy some standard as benchmark, whether of general rationality (Dennett) or of proximity to one's own psychological makeup (Davidson). It is by reference to this benchmark that the representor can zero in on the best available representation. And importantly, this is where the first-person perspective finds its footing. For as interpretationists often stress, the materials for this benchmark will typically come from the representor's *own outlook* – whether her own beliefs, desires, intentions, etc.; or her own (as she sees it) rational take on the situation. In either case, this will involve the representor putting herself in the shoes of the represented. In representing the target agent as inferring some conclusion-belief or as forming some intention, the representor's is asking, 'What does she believe?', 'Which goal is she aiming for?', 'What does she desire?', etc. And answering those

[14] See for example, Baker (2013), Ch. 4.
[15] For discussion, see Root (1986).
[16] See Levy (forthcoming) for a critique of neo-Ryleanism about self-knowledge and self-understanding.

questions involves appealing to the representor's benchmark, which results in trading the above set of questions for their surrogates: What is to be believed?, What is to be done? etc., thereby occupying the reasoner's own perspective. In this way, the interpretationist mode of explanation recognizes and finds a place for the first-person point of view of the reasoning agent.[17]

It has been noted above that, contra radical interpretationism, realism about the mental *is* assumed here. Nonetheless, the interpretation of the target-agent is not thought (as it is for example by Lewis [1974: 334]) to directly reflect or determine her mental reality: The representation of the reasoning agent as undergoing a fully conscious and explicit process is, to repeat, often enough an idealized *mis*representation. This is highlighted again when turning to a closely related and potentially instructive analogy with the present proposal, namely *representation theorems* – most famously (though by no means exclusively) deployed within economic theory.

Decision theory or Expected Utility (EU) theory is the classical theory of decision under risk/uncertainty. It is essentially a paradigm formalization of the rational interaction of beliefs and desires within interpretationism.[18] The basic insight of EU theory is that the rational choice maximizes an agent's expected utility, i.e. the sum of her probability-weighted utilities from each of the available outcomes. Importantly for present purposes, proponents of EU theory typically regard the suggestion that to choose rationally is to maximize EU as deriving from a more fundamental source – namely, the agent's preference relation. The latter is assumed to satisfy certain axioms corresponding to rational requirements of consistency, such as transitivity and completeness. An agent with a consistent preference relation is then guaranteed, through a representation theorem, to behave *as if* she were aiming to maximize her expected utility (von Neumann & Morgenstern 1944, Savage 1954). That is, what holds in fact is the agent's preference relation meeting certain formal (rational) constraints, but the representation theorem allows to go

[17] Cf. Moran (2017), Ch. 12.
[18] Two comprehensive treatments of decision theory are Joyce (1999) and Bermudez (2009).

beyond that reality and view the agent's behaviour as equivalent to her having a utility function which she consciously maximizes. No representation theorem is formulated (let alone proved) here, yet a structurally similar move is developed in identifying the conditions under which one can be represented as if undergoing a fully conscious and explicit process of reasoning.[19]

3.3. What does it take to be representable as if reasoning fully consciously and explicitly?

The reader will recall that the proposal of the present section is strictly independent of the foregoing discussion: One may accept that reasoning can be less than fully conscious and explicit and accept also that reasoners are those who are representable as if reasoning consciously and explicitly, and yet opt for plugging in a different set of conditions for their being so representable than the ones proposed below. With that proviso in place, what are the proposed conditions?

First, it seems clear that there must be a 'movement of mind': some process or event of transition between the mental starting point(s) and the conclusion. Otherwise, in line with the present assumption of realism, there simply could not be any process or event *of reasoning* in particular.[20] Second, more controversially, a reasoning agent must, in forming the conclusion-attitude, *respond* to the premise-attitudes, in the way characteristic of reasoning (the notion of 'inferential response' is explicated immediately below). For concreteness, consider again the example given above:

---

[19] Two standard interpretations of decision theory are available, often labelled 'behaviorist' and 'mentalistic'. According to the former, representation theorems demonstrate that a rational agent's *choice behaviour* is equivalent to what it would be were she to maximize a utility function, while the maximization of the function itself is treated as strictly hypothetical. On the mentalistic interpretation, the agent's utility function and its maximization are treated as psychologically real phenomena. Since the present approach denies that a fully conscious and explicit process of reasoning necessarily takes place when one reasons, it is closer in spirit to the behaviorist interpretation in understanding the representation in strictly 'as if' terms. For further discussion of the different interpretations of decision theory, see Dietrich and List (2016), Okasha (2016), and Thoma (forthcoming).

[20] This condition is rejected by some, as already noted (n. 4 above).

(1) If the government's new measures aren't effective, house prices will continue to soar. The government's new measures aren't effective. So, house prices will continue to soar.

If an agent is undergoing a mental process corresponding to the above pattern that involves her responding inferentially to the premise-attitudes by forming the conclusion-attitude, then the claim is that she is genuinely reasoning because she is representable as if undergoing the fully conscious and explicit process. Representing her in this way allows for a straightforward and powerful explanation of her behaviour. It also makes vivid that her reasoning is person level and something she may be held responsible for. Recall Boghossian and others' claim discussed above, that only fully conscious and explicit reasoning is person-level and responsible. On the present view, these philosophers are overgeneralizing from idealized conditions that serve to make it vividly apparent that reasoning is person level and responsible, into thinking that those conditions must be necessary for reasoning to be so. In fact, genuine reasoning does not require consciously believing the premise-attitudes; a reasoning agent may instead have an attitude towards the premises that falls short of believing their content, and/or have attitudes with contents that only approximate the above. Nor must the agent register the *support* provided by the premises for the conclusion. Assuming she is undergoing some mental process, the reasoning agent must additionally only form the conclusion-attitude in response to the premises-attitudes. But what is it to respond in this particular way?

Recall Siegel's (2019) 'Pepperoni' case set out above (§2.2). It illustrates nicely how one can respond to various features without consciously registering them nor how they support the conclusion. Siegel herself highlights the importance of the inferential response to the project of understanding reasoning and does much to illuminate it. Yet she deliberately stops short of defining the notion (2019: 23). That *Pepperoni* involves an inferential response distinguishes it from cases of mere succession of attitudes, where e.g. one forms the conclusion-attitude as a result of

taking a blow to the head. In such cases as the latter, one's forming the attitude does not involve responding to the premise-attitudes in *any* way. But there are plenty other circumstances where one is not reasoning and yet one's forming the conclusion-attitude may well be thought of as a kind of response – for example, if it is formed as a result of conditioning, association, mind-wandering, and so on. It seems, therefore, that having a definition of the particularly inferential response would be handy.

In lieu of a developed account, the next and concluding section will offer some speculative ideas on how we can approach this project by exploring the aims of reasoning itself (not to be confused with the aims of *representing* agents as reasoning consciously and explicitly, discussed earlier). But first, the section recapitulates the structure of the overall argument.

4. CONCLUSION

The overarching argument of this paper comprised of three broad stages. The first stage attacked the predominant conception of reasoning for its specious restriction to fully conscious and explicit episodes. This was done by raising counterexamples and considering (extant and possible) responses, as well as by critically examining the reasons adduced for the restriction by its proponents. Then, a replacement for the faulty schema of reasoning as a fully conscious and explicit process of transition between attitudes was proffered. According to this alternative schema, a reasoning agent is one whose psychology meets certain conditions that allow to represent her as if she were undergoing a fully conscious and explicit process. Understanding reasoning in these terms extends to cover unconscious and implicit episodes while providing a straightforward and powerful explanation of reasoners' behaviour, which moreover highlights their being person-level and apt for attributions of responsibility.

The next step taken in explaining the nature of reasoning was to propose one way to fill in the proposed schema. Minimal conditions for representing agents as reasoning in the fully conscious

and explicit sense were spelled out. (Readers not convinced by this stage of the argument may come up with alternative ways of filling it in.) As already noted, the epilogue to this paper will extend the argument by putting forth some initial ideas on how to elaborate these conditions further. In doing so, it will help to contrast these ideas with related ones contained in a rival account of reasoning discussed above, due to McHugh & Way.

* * *

Recall that according to McHugh & Way's (2018a) account, reasoning is a functionally-defined process that constitutively aims at getting correct, right, or "fitting" attitudes – such as true/knowledgeable beliefs or permissible intentions. Now, calling on the aims of reasoning does seem a promising way of getting clear on the nature of this phenomenon. The problem, however, is that the particular aim of getting fitting attitudes is too broad, which points to a potentially serious flaw in McHugh & Way's account. To see this, consider the following patterns:

? Today is Wednesday. So, 71 is prime.

? Today is Wednesday. So, I shall save this drowning child.

Here, the conclusion-attitude is *necessarily* fitting and so such patterns (trivially) preserve fittingness. But they are not intuitively recognizable as episodes of reasoning. Now McHugh & Way are aware of this last problem (see their 2018b, §5.2.) And while there is no space here to consider their proposed solution, the point stands that absent further qualification or restriction, stating the aim of reasoning as getting fitting attitudes seems overly inclusive. A brief sketch of an alternative proposal will now be offered, on which reasoning is regulated by a different aim, viz. that of settling practical and theoretical inquiries – or for short, the aim of settling questions. This broad aim provides one way of fleshing out the inferential response introduced earlier. Connecting this idea to previous steps, the suggestion is that in order to be represented as reasoning fully

consciously and explicitly, agents must respond inferentially. And responding inferentially in turn involves aiming to settle inquiries, in a sense to be explained now.

A *theoretical* inquiry is an inquiry into some matter or question that one is curious about or is anyway driven to figure out the answer to. This could involve an attempt to understand, predict and/or explain some part of the inanimate or animate (including the human) world. Will house prices continue to soar? Why is that person turning on the light with their head instead of their hand? How long does it take to get from King's Cross to Bethnal Green? A *successful* theoretical inquiry culminates in figuring out the (correct) answer to one's object of inquiry. This will plausibly involve at least either believing truly or knowing the answer (for present purposes we can afford to remain neutral here and avoid this difficult choice). But of course not all reasoning – theoretical or practical – is successful; one would still count as reasoning even if the mental process one is undergoing were *un*successfully regulated by the aim of settling the relevant question (more on this in a moment).

Practical inquiry is similarly directed at settling questions, though not theoretical but practical ones. For example: How can I make pesto sauce? Should I join Extinction Rebellion? How do I go about getting what I need? Now such questions could be posed theoretically, for instance if one is idly entertaining them. What would make the process they govern one of *practical* reasoning in particular is determined by the type of attitude that settles the governing question, or in other words by how the process concludes. A process of practical reasoning aims to settle question(s) by forming a practical conclusion-attitude (typically, an intention) or by acting (see n.1 above) rather than by believing. But what does it take for a process to aim at or be regulated by settling some question, practically or theoretically?

The idea that some objects and phenomena may be fruitfully understood in terms of their essential aims is familiar. It clearly holds for such artefacts as knives and umbrellas, which are designed to execute certain specific functions. But some states and activities arguably also have

definitive aims (Williams 1973, Shah 2003, Velleman 1989, Korsgaard 2009). Playing chess has the aim of checkmating one's opponent, for example, while house building has the aim of providing shelter. Famously, contrasting these respective aims reveals an important difference in how each activity differentially relates to its aim: The latter aim but not the former can be said to be *constitutive* of the activity in question. That is, one may arguably play chess without trying to checkmate one's opponent, by simply moving the pieces around the board for fun, say. But one is simply not engaged in housebuilding unless one is aiming to provide shelter in what one is doing. The claim that reasoning aims at settling questions is meant in the constitutive sense of 'aim at'.

To aim to provide shelter is (in part) to be sensitive to whether what one is doing is in fact promoting this aim or not; one cannot be thought of as building a house if one is indifferent to the roof's imminent collapse. Similarly, one is not reasoning about the way to get from King's Cross to Bethnal Green if one is manifestly indifferent to the fact that one's ultimate plan only gets one as far as Liverpool Street. Having the aim of settling some question *may* involve representing it in one's reasoning. When it does – when one explicitly asks oneself e.g., 'Now how do I get from King's Cross to Bethnal Green?' – one's reasoning will tend to seem more like *deliberation*, understood here as a species of reasoning that is typically conscious and explicit, as well as more open-ended and protracted compared to other instances of the genus. In any event, it is not necessary that a reasoner represent either the question she is aiming to settle or indeed the aim of settling it (which would raise regress worries familiar from Lewis Carrol [1895]). Rather, her having the aim may be reflected in the manner in which her reasoning unfolds.

For example, her reasoning may conform to good or correct patterns, which actually facilitate the aim of settling whether *p* (e.g. modus ponens) or how to go about *V*-ing (e.g. means-ends reasoning). And even if her reasoning only approximates such patterns, she may still be said to be aiming (though unsuccessfully) at settling the question. Indeed, even if the pattern one's reasoning follows badly fails to facilitate the aim – for example, if it's an instance of affirming the consequent

or the gambler's fallacy – it is still possible that one is aiming to settle some question. This would be the case for instance if one is unaware that the pattern is fallacious and would renounce it were one to become aware. If however, one *does* know the pattern is fallacious but is entirely unperturbed by this fact, then it is hard to see one as reasoning (cf. McHugh & Way 2018a: 182-3).

Some further manifestations of one's aiming to settle a question (without representing this aim) do not bear directly on the pattern of one's reasoning. Rather, they involve being responsive to a host of subsidiary questions, such as e.g. whether one is in a position to settle the question or not (perhaps one has not gathered sufficient evidence for one or more of the premises, for instance); or the higher-order question of whether the question to be settled is even worth pursuing, or worth pursuing now rather than later, and so on. In being responsive to such concerns, one is again being responsive to the project of correctly settling one's guiding question. After all, if one comes to view one's guiding question as not worth pursuing, one will tend to 'settle' it by suspending judgment or simply by exhibiting indifference.[21] [22]

Suspending judgment or exhibiting indifference are not ways of settling a question, as they do not involve accepting any answer. But are there limits on the answers that *can* be considered to settle the question animating one's reasoning? We should not demand that the answer be correct, not even approximately; as already noted, reasoning is sometimes bad (and even good reasoning can arguably terminate at the wrong place). But this does not mean that any answer, however

[21] Normative questions of the sort cited in the text, to do with conducting inquiries, figure in an influential recent plea by Jane Friedman (2020; forthcoming) to reorient epistemology beyond what she sees as the overly narrow 'doxastic' conception and towards a more expansive 'zetetic' one, which sees inquiry as a whole as the proper domain of epistemology.

[22] Some readers may balk at the suggestion that reasoning always aims to settle inquiries. Could not an agent who casually notices that some set of propositions is inferentially related be thought of as reasoning? To some extent, I suppose this is matter of conflicting intuitions: to this writer at least, it seems that if it happens to occur to one that since the PTA meeting is scheduled for Wednesday, the PTA meeting will take place on one's birthday, is more like someone who has experienced a flash of insight than a reasoner. However, it might help the unconverted to bear in mind that even some such cases as the one just described would qualify as aiming to settle a question, given the liberal understanding provided in the text of what this activity involves. Thanks to an anonymous reviewer for discussion here.

absurd or unrelated to the question, could be taken to settle it. Here we can build on a familiar observation noted earlier when discussing the interpretationist framework (§3.2), namely that representing an agent as reasoning in a particular way involves imposing some (albeit thin) standard of rationality or proximity to the representor(s) own perspective. The representor appeals to this benchmark in choosing the best available representation. And part of choosing the best representation is identifying answers that could be thought to settle the question for the agent. Now importantly, if the reasoner's answer is not recognizable as potentially settling her question, then she cannot be represented as reasoning consciously and explicitly, and hence on the present approach, she does not count as reasoning at all. We can see an illustration of this by revisiting the counterexamples presented earlier to McHugh & Way's account. Those counterexamples, recall, involved patterns of attitudes that preserve fittingness but intuitively fail to count as reasoning. The present approach delivers the correct verdict on such cases: it rules them out precisely because it is hard to see what question the agent could be aiming to settle in e.g. responding to the belief that today is Wednesday by forming the belief that 71 is prime, or the intention to save a child from drowning.[23]

More could be said to defend the thought that reasoning constitutively aims at settling questions. But hopefully enough has been said already to make the idea workable in future research. It was argued earlier that reasoners are those who can be represented as undergoing a fully conscious and explicit mental process involving a conclusion-attitude that constitutes an *inferential response* to the premise-attitudes. One way of explicating the notion of 'inferential response' sees it as a kind of response given in the service of settling some theoretical or practical question, as adumbrated above. All told, reasoners may be understood as those who respond to their premise-attitudes in this particular way, making them representable as if reasoning fully consciously and explicitly.

[23] I am grateful to an anonymous referee for discussion here.

References

Baker, Lynne Rudder 2013. *Naturalism and the First-Person Perspective*. Oxford University Press.

Bekoff, M., Allen, C. and Burghardt, G. M. (eds.) 2002. *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. MIT Press.

Bermudez, J. L. 2009. *Decision Theory and Rationality*. Oxford: Oxford University Press.

Balcerak-Jackson M. & Balcerak-Jackson B. (eds.) 2019. *Reasoning: New Essays on Theoretical and Practical Thinking*. Oxford University Press.

Boghossian 2019. "Inference, Agency, and Responsibility." In Balcerak-Jackson & Balcerak-Jackson, pp. 101-124.

Boghossian 2014 "What is Inference?" *Philosophical Studies* 169: 1–18.

Broome, J. 2014. "Comments on Boghossian". *Philosophical Studies* 169: 19–25.

Broome, J. 2013. *Rationality Through Reasoning*. Wiley-Blackwell.

Broome, J. 2002. "Practical reasoning", in Bermúdez & Millar (eds.), *Reason and Nature: Essays in the Theory of Rationality,* pp. 85-111. Oxford University Press.

Bueno, O., French, S. & Ladyman, J. 2002. "On Representing the Relationship between the Mathematical and the Empirical", *Philosophy of Science* 69(3): 452–73.

Callender, C. & Cohen, J. 2006. "There is no problem of scientific representation", *Theoria* 21: 67–85.

Carroll, L. 1895. "What the tortoise said to Achilles", *Mind* 4: 278–80.

Contessa, G. (2007), "Scientific representation, denotation and surrogative reasoning", *Philosophy of Science* 74: 48–68.

Da Costa, C. A. & French S, 1990. "The Model-Theoretic Approach to the Philosophy of Science", *Philosophy of Science* 57(2): 248–65.

Dancy, J., 2018. *Practical Shape: A Theory of Practical Reasoning*, Oxford University Press.

Davidson, D. 1983. "A Coherence Theory of Truth and Knowledge." In E. LePore (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, pp. 307-319.

Davidson, D. 1973. "Radical Interpretation", *Dialectica* 27 (1): 314-328.

Dennett, D.C. 2009. "Intentional Systems Theory". In B. McLaughlin, A. Beckermann & S. Walter (eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford University Press.

Dennett, D.C. 1987. *The Intentional Stance*. MIT Press.

Dietrich, F. and List, C. 2016. "Mentalism versus behaviourism in economics: a philosophy-of-science approach". *Economics and Philosophy* 32*:* 249-281.

Fernandez, P., 2016, "Practical Reasoning: Where The Action Is." *Ethics* 126: 869-900.

Friedman, J. 2020. "The Epistemic and the Zetetic". *Philosophical Review* 129: 501-536.

Friedman, J. (forthcoming). "Zetetic Epistemology". In *Towards an Expansive Epistemology: Norms, Action, and the Social Sphere*. Routledge.

Giere, R.N. 2004. "How Models Are Used to Represent Reality", *Philosophy of Science* 71(5): 742–52.

Gergely, G., Bekkering, H. & Király, I. 2002. "Rational imitation in preverbal infants", *Nature* 415: 755.

Hieronymi, P. 2008. "Responsibility for Believing", *Synthese* 161: 357–373.

Hlöbil, Ulf (2014). "Against Boghossian, Wright and Broome on inference." *Philosophical Studies* 167(2): 419-429.

Hornsby, J. 2012. "Actions and Activity". *Philosophical Issues*, 22, pp. 233–45.

Hurley, S. & Nudds, M. (eds.) (2006). *Rational Animals?* Oxford University Press.

Jenkins, D. 2021. "Reasoning and its Limits", *Synthese* 199: 9479–9495.

Joyce, J. 1999. *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Korsgaard, C. M. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford University Press.

Levy, Y. Forthcoming. Neo Ryleanism about Self Understanding. *Inquiry*.

Levy, Y 2020. "Events, Processes, and the Time of a Killing". *Ratio* 33: 138-144.

Levy, Y. 2019. 'What is "Mental Action"?' *Philosophical Psychology* 32: 971-993.

Lewis, David K. 1974. "Radical Interpretation", *Synthese* 27 (3-4): 331-344.

Lurz, R. 2009. *The Philosophy of Animal Minds: New Essays on Animal Thought and Consciousness*. (Cambridge University Press).

Marcus, E. 2021. *Belief, Inference, and the Self-Conscious Mind*. (Oxford University Press).

McHugh, C. & Way, J. 2018a. "What is Reasoning?", *Mind* 127: 167-196.

McHugh, Conor & Way, J. 2018b. "What is Good Reasoning?" *Philosophy and Phenomenological Research* 96: 153-174.

McHugh, C. & Way, J. 2016. "Against the Taking Condition." *Philosophical Issues* 26: 314-331.

Moran, R. 2017. *The Philosophical Imagination: Selected Essays*. Oxford University Press.

Neta, R. 2013. "What is an inference?" *Philosophical Issues* 23: 388–407.

Nichols, S. & Stich, S. 2003. *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press.

O'Brien, L. & Soteriou, M. (eds.) 2009. *Mental Actions* (Oxford University Press).

Okasha, S. 2016. "On the Interpretation of Decision Theory", *Economics and Philosophy* 32**:** 409–433.

Pautz, A. 2013. "Does Phenomenology Ground Mental Content?", In U. Kriegel (ed.), *Phenomenal Intentionality*. Oxford University Press, pp. 194-234.

Root, M. 1986. "Davidson and Social Science". In LePore, E. (ed.), 1986. *Truth and interpretation*: *perspectives on the philosophy of Donald Davidson* (John Wiley & Sons): 273-307.

Savage, L. J. 1954. *The Foundations of Statistics*. New York, NY: Dover.

Scanlon, T. 2008. *Moral Dimensions: Permissibility, Meaning, and Blame*. Harvard University Press

Siegel, S. 2019. "Inference Without Reckoning", in Balcerak-Jackson & Balcerak-Jackson.

Shah, N. 2003. "How truth governs belief", *Philosophical Review* 112: 447-482.

Smith, A.M. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life", *Ethics* 115: 236–271.

Steward, H. 2013. "Processes, Continuants, and Individuals," *Mind* 122: 781-812.

Stout, R. (ed.) 2018. *Process, Action, and Experience*. Oxford: Oxford University Press.

Strawson, G. 2003 'Mental Ballistics or the Involuntariness of Spontaneity', *Proceedings of the Aristotelian Society* 103: 227–257.

Suárez, M. (2004). "An inferential conception of scientific representation", *Philosophy of Science* 71: 767–79.

Teller, P. 2001, "Twilight of the Perfect Model Model", *Erkenntnis* 55(3): 393–415.

Tenenbaum, S., 2007. "The Conclusion of Practical Reason" in S. Tenenbaum (ed.) *New Trends in Philosophy: Moral Psychology* (Rodopi), pp. 323-343.

Thoma, J., forthcoming. "Folk Psychology and the Interpretation of Decision Theory." *Ergo*.

Tucker, C. 2012. "Movin' On Up: Higher-Level Requirements and Inferential Justification." *Philosophical Studies* 157: 323–340.

Tye, M. 2016. *Tense Bees and Shell-Shocked Crabs: Are Animals Conscious?* (Oxford University Press)

Valaris, M. 2017. "What reasoning might be". *Synthese* 194: 2007-2024.

Velleman, D. 1989. *Practical Reflection*. Princeton University Press.

von Neumann, J. & Morgenstern, O. 1944. *Theory of Games and Economic Behavior*. Princeton University Press.

Weir, A., Chappell, J. & Kacelnik, A. 2002. "Shaping of hooks in New Caledonian crows", *Science* 297: 981.

Williams, B. 1973. "Deciding to believe". In his *Problems of the Self* (Cambridge University Press): 136–51.

Williams, J.R.G. 2020. *The Metaphysics of Representation*. Oxford: Oxford University Press.

Wright, C. 2014. "Comment on Paul Boghossian, 'What is inference?'" *Philosophical Studies* 169: 27-37.