

Synthese: Springer Online May 2014. The final publication is available at Springer via <http://dx.doi.org/10.1007/s11229-014-0487-1>.

In My 'Mind's Eye': Introspectionism, Detectivism, and the Basis of Authoritative Self-Knowledge

Cynthia Macdonald

Abstract

It is widely accepted that knowledge of certain of one's own mental states is authoritative in being epistemically more secure than knowledge of the mental states of others, and theories of self-knowledge have largely appealed to one or the other of two sources to explain this special epistemic status. The first, 'detectivist', position, appeals to an inner perception-like basis, whereas the second, 'constitutivist', one, appeals to the view that the special security awarded to certain self-knowledge is a conceptual matter. I argue that there is a fundamental class of cases of authoritative self-knowledge, ones in which subjects are consciously thinking about their current, conscious intentional states, that is best accounted for in terms of a theory that is, broadly speaking, introspectionist and detectivist. The position developed has an intuitive plausibility that has inspired many who work in the Cartesian tradition, and the potential to yield a single treatment of the basis of authoritative self-knowledge for both intentional states and sensation states.

It is widely accepted that knowledge of certain of one's own mental states is authoritative in being epistemically more secure than knowledge of the mental states of others, despite the fact that such knowledge seems typically to be neither inference-based nor empirically evidence-based. Given that beliefs that are not based on evidence are not generally thought to be more reliable than ones that are, theories of self-knowledge have largely appealed to one or the other of two sources to explain the special status that attaches to at least some of it. The first, 'detectivist', position, appeals to a basis for self-knowledge which is perception-like in involving a reliable causal mechanism that enables subjects 'detect' their own mental states, whereas the

second, ‘constitutivist’, one, appeals to the view that the special security awarded to certain self-knowledge is a conceptual matter.¹

In what follows, I argue that there is a fundamental class of cases of authoritative self-knowledge, ones in which subjects are consciously thinking about their current, conscious intentional states, that is best accounted for in terms of a theory that is, broadly speaking, introspectionist and detectivist. It is detectivist in appealing to a reliable causal mechanism between two states, the reflecting state and the state reflected upon, that enables subjects to become aware of their own mental states. It is introspectionist rather than extrospectionist in appealing to an inner process of awareness rather than appealing to perception of things and phenomena external to the mind as a basis for finding out about one’s own mental states. The theory developed departs from other treatments of the authoritative status of self-knowledge in two quite different ways. First, it is not, like most other introspectionist theories are, a version of an inner sense view in that it does not appeal to a ‘self-scanning’ or ‘self-monitoring’ mechanism (cf. Armstrong 1981, Lycan 1996).² Second, it differs from other recent positions that reject introspectionist theories altogether. These include both neo-expressivist views of the kind advanced by Bar-On (2004, 2009) and positions like that endorsed by Peacocke (2009), which reject constitutivism but do not endorse introspectionism.

My overall aim is to set out a version of a detectivist position that will provide a satisfactory response to the particular problems encountered by a Burge-type ‘same-

¹ The terminology of ‘detective’ and ‘constitutive’ is used by Wright in his 2001 discussion of self-knowledge, and by Finkelstein (2008), though the use of ‘constitutive’ to mark out the constitutivist position is common to many other discussions of self-knowledge (see, for example, Burge 1996, Bilgrami 2006, Bar-On 2004, and Shoemaker 2009). Alternative terminology is used, for example, by Bar-On 2004, and by Zimmerman 2008.

² So the theory I am proposing differs both from what Finkelstein (2008) calls ‘old detectivism’ (which takes the sort of detection that yields self-knowledge to be infallible rather like the way Russell (1912) construed subjects’ acquaintance with sense-data) and from what he calls ‘new detectivism’ (which construes self-awareness to be a kind of ordinary perception in the way that e.g., Armstrong 1981 does). Section 2 explains in what ways introspection is and is not like perception.

order' view, which holds that in cases of 'basic' self-knowledge (the so-called *cogito*-like cases) the state known is literally a part or constituent of the knowing state (Burge 1996, 1998).³ While I do not deny the existence of such cases, by far the majority of situations in which we have authoritative self-knowledge do not fall into this category, and at the core of these, I will argue, lie cases of current conscious thoughts about one's thoughts, when they count as knowledge. The position I set out has an intuitive plausibility that has inspired many who work in the Cartesian tradition. It has come to be considered deeply flawed, but many of the points made against certain versions of Cartesianism or particular Cartesian claims (for example, that subjects have immediate epistemic access only to an inner world, which acts as an intermediary between them and the external world) can be accommodated by a subtler version of an introspectionist, detectivist view.

The remainder of this paper falls into five sections. Section 1 sets out the Burgean same-order view and points out the difficulties it has accounting for a number of cases of current conscious thinking, ones where it is intuitively plausible to suppose that there are two distinct states, the state thought about and the state of thinking about it, which are wholly distinct from and contingently causally related to one another. These cases help to motivate detectivism, which is a two-order rather than a same-order view. Common to them is a certain distinctive kind of phenomenology, the presence of which plays an important role in motivating introspectionist, detectivist positions but which plays no role in motivating the Burgean explanation of our entitlement to self-knowledge. It then specifies a plausible constraint on an adequate

³ Ned Block (2007) uses the term 'same-order' to refer to the type of constitutivist view held by Burge (1996, 1998) and others (Shoemaker 1996, 2009), according to which first-order thought is literally a constituent of the reflective thought about it (Burge's paradigmatic examples being the so-called *cogito*-like cases).

account of these other types of cases - a phenomenological one - similar to one that is often reasonably taken to constrain theories of the nature of secondary qualities.

Sections 2 through 4 set out to meet this constraint in two stages. The first (Section 2) develops an introspectionist account of the nature of the special epistemic security that attaches to certain self-knowledge – an account of in what that special security consists. This fleshes out an intuitive sense, in which subjects' knowledge of their own thoughts is epistemically direct and immediate in the relevant cases, which forms a necessary part of the account of the special epistemic security that attaches to it. This characterisation provides an alternative to the Burgean one which is compatible with the detectivist view in being compatible with a two-order rather than a same-order view of the relation between the reflecting states and the states reflected upon.

The second stage (Sections 3 and 4) accounts for the sense in which subjects' second-order states are *about* their first-order ones in cases of authoritative self-knowledge in a way that vindicates a version of the detectivist view by addressing two problems: a temporal one, and a semantic one. The temporal problem is how it is that one can both think and think about a thought at the same time without the former being a constituent or part of the latter. The semantic problem is how reference to the first-order thought is secured. Section 5 concludes by articulating the combined introspectionist, second-order position and explaining how authority is secured.

Although I argue that a second-order view that takes the relation between one's first-order states and one's second-order ones in the relevant cases to be contingent and causal rather than constitutive is all that is needed to meet the phenomenological constraint in the relevant core cases, I do not discuss at length the many versions of constitutivism. Since there are substantial and important differences between these

versions, this would require a separate work devoted entirely to that topic. The strategy, rather, is to articulate the best version of a type of position that has inspired many in the Cartesian tradition, one that has a great deal of intuitive plausibility and merits serious scrutiny. That it does have such intuitive plausibility argues in its favour; the onus is therefore on opponents of the position to argue against it. The core constitutivist thesis against which I defend the view is that the knowing second-order thoughts must be viewed as constitutively related to the known first-order ones, either by being partly constituted by those first-order thoughts, or by being non-contingently (even if causally) related to them, both of these ruling out a merely contingent causally reliable source of epistemic entitlement.

I restrict my attention in the bulk of what follows to the cases of current conscious thinkings, musings, wonderings, and the like - intentional states that lie at the core of more epistemically committed intentional states such as current conscious beliefs, and so also at the core of intentional states like standing beliefs. It is states like these whose special epistemic status is most fundamental to account for, since it is hardly credible that one can know what it is one currently believes without knowing what it is one currently thinks. Authoritative knowledge of such conscious states is knowledge that can be expressed in conscious thoughts of the form 'I am currently consciously thinking (musing, wondering, etc.) that p ', for some propositional content p . The concern here is specifically with the limited task of explaining authority with respect to knowledge of the contents of one's current conscious thoughts (that it is p that one is thinking).

Section 1: Motivating the Second-Order, Detectivist View

Constitutivism about our cognitive awareness of our own mental states is the view that there is a non-contingent, conceptual connection between certain of our so-called first-order mental states and our thoughts, desires, or beliefs about them. Although there are many versions of the view (cf. Bilgrami 2006, Wright 1998, Boghossian 1989) as well as ‘hybrid’ positions that are constitutivist with respect to certain types of mental states but not with respect to others (cf. Boyle 2009, Zimmerman 2008), I want to focus attention on a version which holds that, in the paradigmatic core cases, the state introspected is literally a part or constituent of the introspective one and is conceptually related to it for this reason (Burge 1996, 1998; Shoemaker 1996, 2009). Burge’s paradigm example is the so-called *cogito*-like cases, which are self-referential and thereby contextually self-verifying for the reason that, in thinking the thought, one makes the thought true; one is, in the very act of thinking that one is thinking that *p*, thinking that *p*. The knowledge possessed by subjects in such cases is infallible for this reason; one is exempt from error when one is truly thinking that one is thinking that *p*. Many constitutivists commit to some restricted version of an infallibility thesis for at least some types of self-knowledge (e.g., Parent 2007, Bilgrami 2010).

I do not contest the view that there are *cogito*-like cases, nor that these form an important, though very limited, class of cases of authoritative self-knowledge. However, they are not best seen as the paradigm cases of authoritative self-knowledge. The main reason is that there are many cases of current, conscious thinking, musing, wondering, and the like, where, in contrast with the *cogito*-like ones, subjects are authoritative in the sense of being in an epistemically more secure position than others with respect to knowing that they are in such states, but they are not infallible. Such situations are ones in which the reviewing thoughts are plausibly viewed as being wholly distinct from and contingently causally related to the thoughts reviewed.

Arguably these constitute by far the majority of cases in which subjects have authoritative self-knowledge.

Burge does not deny that there are situations in which the reviewing thoughts are distinct from and causally related to the thoughts reviewed. Nor does he deny that the reliability of these causal relations forms a necessary supplemental part of the conditions that must be met for the reviewing thoughts to qualify as knowledge. However, he maintains that the special epistemic entitlement to self-knowledge and to the authority subjects have with respect to knowing the contents of the thoughts reflected upon in such cases has a source other than the reliability of the causal relations that hold between them. On his view, the basis for this special epistemic status lies in the role that self-knowledge plays in constituting the rational agent, and this role is a constitutive one, not a merely causally reliable one. The entitlement consists simply in functioning in an appropriate way in accordance with norms of rationality or reason, and it has two sources. One is the special role that such knowledge plays in critical reason and reasoning. The other is the special (“constitutive”) relation between judgements involving self-knowledge and their subject matter, or truth, or between one’s judgements about one’s thoughts and those judgements’ being true. Here, in brief outline, is the argument (Burge 1996).

Critical reasoning involves critically assessing, evaluating, adjusting and correcting one’s putative reasons *as* reasons. It involves the ability to employ reason and reasoning to assess and criticise the reasonableness of reasoning itself. So critical reasoning is reasoning that is under the rational *control* of the reasoner. One cannot bring under one’s control something of which one is unaware. In order for reasons to be under the rational control of the reasoner, then, she must normally be knowledgeable about her thoughts. She must know the contents, but also the attitudes,

since critical reasoning is not just concerned with evaluating propositional connections, but also concerned with confirming, adjusting, and correcting attitudes.

So control over one's reasoning requires knowledge of one's attitudes and contents. But it also requires that the relation between one's judgements about one's thoughts and those judgements' being true be constitutive: it cannot be one that could systematically fail to hold. If it could, then reflection could add nothing to the reasonableness of critical reasoning. But reflection *does* add something: it brings one's attitudes under the rational control of the thinker. This constitutive, or non-contingent, relation between the subject matter and the truth of one's judgements about one's thoughts is the second source of subjects' entitlement to self-knowledge.

In order to illustrate why this account of our entitlement to self-knowledge is not the natural interpretation of the kinds of situations I have in mind, let me first present two such types of case, both of which involve current, conscious thinking, and then explain why such cases have features that present problems for the Burge-type view.

First, there are situations involving occurrent propositional memory, where one comes to think, and partly on that basis to judge, that one remembers that something is the case on the basis of a current conscious propositional memory state (which may or may not be veridical). Suppose that Sam asks Joe what opera he wants to see. It seems to Joe that there is an opera he has wanted to see, but nothing immediately comes to mind. He tries to remember, and after thinking for a moment or two, it occurs to him that he wants to see *Aida*. Its occurring to Joe that he wants to see *Aida* is a subjective, conscious recollection, a state that is engaging his conscious attention. It appears to Joe as though he is remembering something and not imagining it, and on this basis he comes to reflectively judge that he remembers that he wants to see *Aida*.

There seem to be at least three states involved in this process. First, there is the original desire to see *Aida*. Then there is the recollection of the desire, the state whose occurrence is currently engaging Joe's attention and on which he is reflecting. Finally, there is the conscious reflective judgement itself, Joe's judging that he remembers that he wants to see *Aida*. In this situation, it is plausible to say that Joe's reason for judging that he remembers that he wants to see *Aida* is his current recollection. The basis for his judging that he remembers is not that his friend told him that he once said that he wanted to see *Aida*, and he trusts his friend's testimony. His basis comes from within, not from without. Since a reason for thinking that one remembers is distinct from what it is a reason for, the states seem to be genuinely distinct from one another and contingently causally connected (cf. Peacocke 1999, who cites different cases of the same kind).⁴

A second type of case concerns one's own thoughts about one's current conscious thinkings, musings, wonderings, speculations, and the like. I might be contemplating the theatre booking I have just made, or wondering whether Cologne is situated on the Seine, or speculating about the chances of my winning my next tennis match - and suddenly realize both that I am doing so and what the content is of my musing, wondering, or speculation. Perhaps I'm driving home after a long day at work, and as I negotiate a familiar junction I realize that I'm not really concentrating on the road at all, I'm actually thinking thoughts about the interview for promotion which I underwent earlier today. I realize that I'm ruminating, thinking about what I thought, said and did at the interview. I suddenly become reflectively aware of my own

⁴ Whether this constitutes an example of introspective knowledge depends on whether there are factive mental states like knowledge ('remembers that' being factive), which can be known directly, i.e., non-inferentially. Here I side with Williamson (1995, 2000), who claims that factive states like knowledge are indeed mental states, which can be known in an epistemically immediate, non-inferential way, rather than with his opponents.

current conscious thoughts that are taking my attention from the state of the road, and also what the contents of those current thoughts are. Here again, there seem to be at least three states involved in the process. First, there are the thoughts and conduct that occurred at the interview itself. Then there are the thoughts – ruminations – that are currently distracting me from the state of the road. Finally, there is the sudden realization, the reflective awareness, of those thoughts.

In both of these types of situations – ones involving thoughts about one's current recollections, and ones involving thoughts about one's current conscious thinkings and the like — it seems that there is a current state that is present antecedently to the sudden realization, and that that state *attracts* one's attention in a direct and immediate way and by this means engages one's thought. In both such types of situations, it seems there are two distinct states that are causally related, the first of which comes to engage and occupy one's attention in thought directly and immediately; this is how the phenomenology of conscious thoughts about one's current recollections, musings, wonderings, speculations and the like presents itself to subjects who undergo them.

What is distinctive about these types of situations is that the first-order thoughts involved in them *present* themselves to their subjects as being distinct from and present antecedently to thoughts about them, and as attracting and engaging the attention of their subjects without mediation by other conscious states and without inference from other conscious states. In other words, the phenomenology of self-knowledge in such situations to the subjects in them is that of second-order, not same-order, thinking, in which subjects' antecedently present thoughts come to engage and occupy their attention.

Although the Burgean view does not deny the existence of two-order cases, it is committed to maintaining that the source of epistemic specialness that attaches to the

self-knowledge they manifest lies in the constitutive, rational connection that holds between reviewing thoughts and thoughts reviewed. The picture that this type of position presents is one whereby our entitlement to self-knowledge flows from our natures as critical reasoners, and the *cogito*-like cases of authoritative self-knowledge are paradigmatic examples of ones where the constitutive, rational connection between a subject's thoughts about her thoughts and the subject matter of those thoughts cannot come apart. It is this that lies at the heart of Burge's view that, unlike perception, there cannot be systematic brute error in self-knowledge, 'accidental' error that does not imply a failure of rationality on the part of the subject.

The types of situations presented here, however, pose problems for this view. They do so because in them the first-order thoughts on which the subject reflects seem *neither* to be *cogito*-like *nor* to depend on the constitutive, rational connection to the reviewing ones that flows from our natures as critical reasoners. In these situations the subject simply *happens* to become aware of, by being attentively attracted to, one's antecedently present conscious thought in an epistemically direct way, i.e., without mediation by or inference from further conscious states. So, such thoughts do not seem to be *cogito*-like (if they were, one could not fail to become aware of the thought that was literally a constituent of one's conscious thought about it). Partly for this reason, and partly because such situations concern states of mere conscious thinking that are very common and – importantly – much less epistemically committed than states of conscious belief, in them it seems possible that one might simply err without impugning rationality; one might simply misidentify the content of one's recollection, or of one's wondering.

Burge (1999: 33) acknowledges that cases like the *Aida* one are non-*cogito*-like, and that in such cases error is possible. His view, nevertheless, is that the warrant or

entitlement to self-knowledge in such cases is special because of the role that they play in a system of critical reasoning, which rules out the possibility of brute error. While this may be plausible for highly epistemically committed states such as beliefs and judgements based upon them, where norms of rationality may reasonably be expected to place a burden on subjects to critically evaluate, retract or amend one's states in the light of other beliefs one holds in accordance with the norms of reason, the kinds of situations cited here concern mere thoughts about one's current conscious recollections, wonderings, musings. Such situations give reason to relegate critical reasoning to a stage further along the road of rational attainment, one that depends on and is supported by cases in which there are contingent causal connections between one's current conscious thoughts and knowledge of them, where such knowledge is secured by introspection, and where the source of epistemic entitlement is the reliability of these causal relations.

Earlier in this section I mentioned that in both of the types of situations cited there is a distinctive phenomenology of thinking and thinking about. The way in which I come to think the second-order thought is due to the way in which the first-order thought attracts and engages my conscious attention. Put another way, in situations of the kind cited the first-order thought contributes to what it is like for a subject to be thinking *about* it.⁵

⁵ That the sorts of cases cited here have a distinctive phenomenology is not to say that the phenomenology is experiential, in the sense in which, for example, the phenomenology of visual or auditory perception, or awareness of pain, is experiential. I expand on this in Section 2. The point is also made by Peacocke:

Perceptual experiences and sensations, on the one hand, and so-called "occurrent" conscious propositional attitudes, on the other, differ in many respects. But there is one property they share. They both contribute to what, subjectively, it is like for the person who enjoys them. A person may try to recall who was Prime Minister of Czechoslovakia when the Soviet Union invaded. It then occurs to this person that Dubcek was the Prime Minister. Its so occurring to him contributes to the specification of what it's like for the person then. It would be subjectively different for the person if it occurred to him (falsely) that it was Husak; and subjectively different again if nothing comes to mind about who was Prime Minister. (1998: 64)

Phenomenological constraints on philosophical theories in areas other than that of self-knowledge are often considered to be both reasonable and desirable. Consider O'Callaghan and Nudds, discussing theories of vision and audition:

Theories of vision aim in part to explain why things look the way they do... Theories of audition similarly aim to explain why things sound the way they do. That includes explaining how audition presents things as being and why auditory experience is organized as it is... The data for these theories thus are partly introspective—they include first-person descriptions of what can be seen or heard, and the way those things look or sound. Reports of phenomenology are data that must be explained by a psychological theory, even if only part of the explanation is that experiences have features accessible to and reportable by the subject (2009: 19)

Here they speak of psychological theories, but O'Callaghan applies the same reasoning in his discussion of constraints on an adequate theory of the nature of sounds:

From the outset, two initial kinds of constraint bear on the theory of sounds. The first is phenomenological. Given that sounds are among the things we hear, how we hear them to be is relevant, *prima facie*, to theorizing about what sounds are. All else equal, an account that captures the phenomenology of auditory perception is preferable to one that does not. (O'Callaghan 2007: 14)

This appeal to phenomenological constraints is not uncommon: the number of theorists about the nature of colours and sounds who appeal to such constraints is legion, and it also surfaces in certain areas of moral theorizing.⁶ Discussions of self-knowledge are also permeated by phenomenological considerations, not only with respect to sensations, but also with respect to conscious intentional states (cf. Peacocke 1999, 2009, Coliva 2009, Kuvlicki 2010, Zimmermann 2008). So the appeal to

He extends the point to ordinary thinking more generally.

⁶ See, for example, David McNaughton, discussing moral realism:

The realist maintains that we should take the nature of our moral experience seriously.... Moral value is presented as something independent of our beliefs or feelings about it; something which may require careful thought or attention to be discovered. The appeal to the nature of our moral experience, to what we might call the *moral phenomenology*, represents the starting point for an argument... (McNaughton 1988: 40)

phenomenological constraints is by no means restricted to theories of sensory experience; it extends into the domain of conscious thinking, musing and the like itself, and appeal to such constraints is not inappropriate in theorizing about the states in this domain.

In the light of this, I propose to take phenomenological considerations of the kind I have indicated in the examples above to constitute a *prima facie* constraint on an adequate theory of authoritative self-knowledge. Specifically, I shall take the following considerations to bear on an account of the nature of authoritative self-knowledge for the kinds of cases that are my primary focus here:

- (PC) Subjects' conscious thoughts about their current, conscious intentional states typically present those states to their subjects:
- (a) in a distinctive epistemically direct and immediate way (i.e., subjects aren't aware of any mediation by or inference from any other conscious state), and
 - (b) as states that they are thinking *about* that are independent of and present antecedently to thoughts about them, attracting and engaging their conscious attention.⁷

In the remaining sections, I develop an account of authoritative self-knowledge that meets (PC) by first addressing part (a), giving an introspectionist characterisation of

⁷ I take this phenomenological constraint to capture what is common to the two types of situations set out in this section – that the subject's knowledge of her own current conscious condition comes to be only with her reflective thought. These are both cases of particular reflective thoughts whose coming to be is prompted by particular conscious occurrent propositional states. This is not to say that the subjects of such reflective thoughts do not have other background dispositional states, standing beliefs and knowledge, which can affect conscious knowledge of their current thinkings. But those standing states are not the ones that attract and engage conscious attention in reflection, and they are not the basis for and do not trigger the reflective thoughts. This is particularly clear in the types of situations that form the basis of the second type of example cited in the present section, thoughts about one's conscious musings, conjectures, speculations, and wonderings, since these are not manifestations of standing knowledge. Further, it may be that a state that is committed to passive short- or long-term memory only comes to be known to a subject by becoming a conscious occurrent memory on which that subject reflects.

the nature of the special epistemic security that attaches to certain self-knowledge that is consistent with a detectivist view (Section 2), and then by addressing part (b) (Sections 3 and 4).⁸

Section 2: Epistemic Immediacy and the Observational Analogy

What is it to have authoritative awareness of one's first-order, conscious wonderings, musings, and the like – in what does such awareness consist? When I suddenly realize that I am wondering whether I got the job, or it strikes me that I remember that the opera I want to see is *Aida*, the conscious thoughts – wondering whether I got the job, remembering that the opera I want to see is *Aida* – seem to come to my attention without mediation by other intentional states, and without inference from any behavioural or other evidence.⁹ This is not unlike the way in which items in my visual

⁸ It might seem to someone sympathetic to Burgean constitutivism that a more charitably-interpreted version of the 'same-order' view is consistent with (PC), as one anonymous referee has suggested. One might concede, for example, that the state reflected upon does precede and is distinct from the reflecting one. But still, one might claim, this leaves open the possibility of meeting PC by maintaining that the previously existing thought becomes, *at the moment of reflecting*, a constituent of the reflecting state. That is to say, one and the same *token* state, with its content, is first thought at one time *t* and then, at a later time *t'*, in reflection, becomes part of the reflecting thought. However, it is difficult to see how this suggestion can meet (PC) if there is just one token state and content. The problem here is a temporal one. In order for that one state and content to exist at two times, *t* and *t'*, it would need to be construed as 'spread out' over and during those times – that is, as existing wholly at *t*, independently of the reflecting one, and also, at *t'*, as existing wholly as part of the reflecting thought. But on that understanding, the whole of the token state could not be plausibly understood as accommodating the phenomenological fact that the state reflected upon at *t'* seems independent of and antecedent to the reflecting one. For, at the moment of reflection, the content embedded in that state is part of the very reflective thought itself and presents itself as such. That very content is, and presents itself as, the content of what one is thinking at *t'*, not as the content of a state that is present independently of and antecedently to *t'*. Put another way, there is nothing in the reflecting state to indicate that that the (token) content of that state is the (token) content of a previously and independently existing state. Rather, it seems that, in order to capture the phenomenology of *both* thinking and thinking about in the non-*cogito*-like cases in a way that meets (PC), one needs to suppose that there are two token thought contents, not just one – one at *t*, and another at *t'*, one of which one is the content of the first-order thinking, the other of which one is part of the content of the reflective state. I develop this thought in section 3, and return to the issue of why two token contents are necessary in connection with the discussion of the paratactic analysis of indirect discourse in section 4.

⁹ This claim is compatible with the view, advanced by Williamson (2000) and others, that when one knows that one is in pain introspectively, the fact that one is in pain is one's evidence that one is in pain; this is the limiting case of *p*'s being evidence on which one believes *p*. The issue here is terminological.

environment, specifically the colours and shapes of perceptible objects, attract my attention in perception. I see a round, green lime on the table before me. My awareness of the greenness, and roundness, of the lime visually present comes to my attention in a way that seems unmediated by other mental states and without inference from behavioural or other evidence. I simply see the green round lime; I see its greenness and roundness.¹⁰ My awareness seems epistemically direct and immediate, and this is part of the phenomenology of perceptual experience of primary and secondary qualities of observable objects.¹¹

We can call this kind of directness *epistemic* directness, and the awareness associated with it - awareness that is non-evidence based, non-inferential, and unmediated by some further intentional (or sensational) state - *direct epistemic access* (cf. Alston 1971). Elsewhere I have argued that there are certain features of observable properties of objects of perception to which subjects have direct epistemic access that mark them off from other sorts of properties of such objects to which subjects do not typically have direct epistemic access.¹² One is that these properties are epistemically *basic* in that they are the favoured and primary means by which knowledge of the objects that have them is obtained. That is to say, typically, in getting to know what an

When I know that I am wondering whether *p* via introspection, my knowledge that I am wondering whether *p* is not based on any evidence *distinct* from the fact that I am wondering whether *p*.

¹⁰ Seeing the colour (in this case, green) just is one's experience of it (on the condition that the subject's experience is veridical and not deviantly caused by the object's having that colour), rather than one's belief that there is a green thing before one. My claim is that perception of the greenness of the lime is no more epistemically mediated by a further mental state than is conscious thinking of what I am thinking via introspection in cases of current, conscious thinking. There is a structural parallel between the two types of case, even though introspecting is not perceiving. (Cf. Peacocke (2009:200), who takes there to be a "significant parallelism of abstract structure" between action awareness (where action awareness includes awareness of *mental* actions such as conscious judgments and calculations) and perceptual awareness, while acknowledging that action awareness is not perceptual.)

¹¹ This is true on at least some accounts of perception. I should note here, however, that there are alternative views of perception according to which my knowledge of the greenness of the lime is not epistemically direct.

¹² See Macdonald 1998, 2007.

object is, these are the first of its properties that one gets to know that it has. Another is that such properties typically *are* as they appear to be to normal subjects in normal conditions; that is, how these properties seem to such subjects to be is typically how they are. This is a version of a kind of transparency claim that many appeal to as a phenomenological constraint on a theory of nature of colours (e.g., Campbell 1994, Boghossian and Velleman 1991, Johnston 1992, Russell 1912); others appeal to it when discussing the nature of audition (cf. O’Callaghan 2007).¹³

These two epistemic features are possessed by properties that fall into the broad category of observable ones, specifically, the so-called primary qualities such as being round, and secondary qualities, such as being green, of objects of perception. My knowledge that the lime visually present before me is green, and that it is round, is plausibly understood as being direct in having these two features. One explanation for this directness is that this instance of the property is presented to me *as* an instance of that property through my sense of sight, and so no evidence or inference is needed to come to know that it is an instance of that property.

This contrasts sharply with other properties of observable objects. The lime visually present before me, for example, instances a certain chemical microstructural property, but this instance is not manifested to me as an instance of that property

¹³ This transparency claim is compatible with at least one interpretation of the thesis, often referred to as ‘transparency’, according to which, in determining or discovering what one believes, one ‘looks through’ one’s belief to a state in the world beyond the mind that that belief is about; and one’s reasons for thinking, judging, or believing that one has a given belief are the very reasons one has for having that belief (Evans 1982). In both cases one looks ‘outward’, toward the world, rather than ‘inward’, as the introspectionist would have it. The interpretation with which the present transparency claim is compatible is known as ‘weak transparency’ (cf. Kind 2003), characterized as follows:

Weak Transparency: it is *difficult* (but not impossible) to attend directly to our experience, i.e., we can most easily attend to our experience by attending to the objects represented by that experience.

It contrasts with a stronger version of the claim, which is incompatible with the view that subjects can have direct epistemic access to certain of their own mental states:

Strong Transparency: it is *impossible* to attend directly to our experience, i.e., we cannot attend to our experience except by attending to the objects represented by that experience.

However, there are compelling objections to this stronger thesis (cf. Gertler 2011, Block 2003).

through one of my senses. So, certain properties are ones to which we have direct epistemic access just because they are observable: whether objects instance them can be determined by unaided observation, though not without conceptualization, of those objects. Not without conceptualization because, in order to have direct epistemic access to a colour property such as the property, green, it's not sufficient that one sees an instance of that property. One must see it *as* an instance of that property.¹⁴ Moreover, the relevant notion of 'seeing as' that is in play is factive, so one cannot see an instance of red *as* an instance of red without seeing that instance. 'Seeing as' is in this respect like 'seeing that', and the 'gap' between 'seeing that *p*' and 'knowing that *p*' is too small to view the former as evidence for the latter, except, perhaps, as a limiting case (cf. n.9).

To say that observable properties are epistemically basic to knowledge of objects that instance them is say no more than that such properties are ones by which objects that instance them are typically known in the first place. Limes are typically known through their determinate colours and shapes. Salt is typically known through its whiteness and saltiness. Equally, to say that observable properties are in general as they appear to be when instances of them are presented to normal perceivers in normal circumstances is to say something about the nature of the properties themselves rather than about the nature of the objects or substances that instance the properties. The nature of gold may be such as to have the chemical constitution Au, but this is compatible with a gold ring's instancing certain observable properties, like goldness and roundness, such that *they* are as they appear to be to normal subjects in normal circumstances.

¹⁴ So the relevant notion of 'seeing' in play here is what Dretske (1969, 1978, 1979, 1995) calls 'epistemic seeing', in contrast with non-epistemic seeing. See also Heil (1983), who develops a similar account of perception, and Audi (2004).

These two features are also possessed by contentful intentional properties of first-order states, properties such as that of being a thought that the lime visually present is green, in the kinds of cases that are here taken to be the core ones. When one thinks of a first-order intentional state while currently consciously thinking it, from the point of view of a second-order intentional state, one's grasp of that first-order state is first and foremost a grasp of it *as* a state of a certain contentful type; the instance is grasped as one of a given contentful property. When reflecting on one's first-order intentional state, this is typically the property by which that thought is known in the first place. Moreover, the contentful properties by which a subject thinks of her current, conscious intentional states are in general as they appear to her to be. Typically, when a current, conscious thought of mine appears to me to be a thought of a certain contentful type, say, a thought that salt is white, it *is* a thought of that contentful type - a thought that salt is white.

To claim this is not to deny that there are significant differences between the way in which these two features apply to observable properties like that of being green or being round, and the way in which they apply to intentional properties, such as the property of being a thought that salt is white. One difference is that there seems to be no experiential dimension to grasp of the latter sorts of properties, whereas there is an experiential dimension to perception of primary and secondary qualities.¹⁵ Another is that observable properties are importantly connected with their effects on normal

¹⁵ But see Horgan and Tienson (2002), and Horgan, Tienson, and Graham (2004) for a view that denies that intentional properties lack an experiential aspect. Here we are not talking about whether there is something it is like to have an attitude toward a content, say, to believe, rather than doubt, it, but about whether there is an experiential aspect to the content of the attitude itself: "Intentional states have a phenomenal character, and this phenomenal character is precisely the what-it-is-like of experiencing a specific propositional-attitude type vis-à-vis a specific intentional content." (Horgan and Tienson 2002: 521). I agree with Shoemaker and others who deny that one experiences the contents of one's current conscious intentional states when one consciously thinks about them in the sense of 'experience' that applies to sensory experience and experience of one's sensations.

perceivers. Observable properties are directly accessible equally in the same, observational way to *all* normal perceivers, whereas the intentional properties of which my thoughts are instances are directly accessible only to me. For my intentional states are directly accessible to me alone; that is at least in part what it is for these states to be my states. Although others may know what type of intentional state I am in by means of observing actions of mine (i.e., in the observational way), only I can know them in ways that others generally do not, and this is what makes for the epistemic asymmetry between my knowledge of my own states and others' knowledge of those states.

Given that there seems to be no experiential dimension to, say, my grasp of the content of my wondering whether Cologne is on the Seine, in what sense is my grasp of such content relevantly like my grasp of the whiteness or saltiness of salt via perception? One way in which they are relevantly alike is this. Just as in visual perception there is a certain distinctive, visual sensory way in which I perceive the whiteness or saltiness of salt – I see the instance of whiteness, or of saltiness, *as* an instance of that property – there is also a distinctive way in which I am aware of the contentful properties of my wondering. I am aware of my wondering *as* a wondering whether Cologne is on the Seine, and that state contributes to what it is subjectively like for me to be consciously aware of it. It would be consciously different for me if, instead, I was wondering whether Cologne is on the Rhine. In neither case is grasp of the instances of the relevant properties mediated by a further mental state, despite the fact that seeing is not wondering, and despite the fact that the epistemic directness of my awareness of the contentful property of my wondering is not in the perceptual, experiential way. Awareness of the contentful properties of one's conscious wonderings need not be by way of perceiving or observing them for it to be

epistemically direct in a parallel way. Those who reject perceptual models of introspection, or more broadly detectivist ones, can and do accept such claims.¹⁶

So I am not recommending a perceptual model of introspection. Rather, I am endorsing certain abstract structural analogies between introspective awareness of the contentful properties of our current conscious thinkings and perceptual awareness of observable properties of objects of perception. With Shoemaker, I agree that there are significant differences between perceptual knowledge and self-knowledge, even in the core cases of knowledge of the contents of one's current, conscious thinkings. A crucial one, already noted, is that my grasp of the contents of such thinkings seems to have no experiential dimension. This difference is due to the fact that, although both in perception and in thought about one's current thinkings there is a certain distinctive way in which properties of objects/states are grasped, the perceptual way is not the introspective way. There is no analogue of perceptual experience, since thinking is not perceiving. One does not experience the contents of one's thoughts, either in the sense in which one experiences objects of perception in vision, or in the sense in which one experiences one's pains. However, this should not prevent us from seeing that, in both cases, the awareness is epistemically direct – noninferential and unmediated by further mental states.

Moreover – and here is a second way in which perception and thought about one's current conscious thinkings are relevantly alike — in both cases epistemic directness is compatible with there being an is/appears distinction. Error is possible because such epistemic directness does not guarantee veridicality. In perception, further conditions, both internal and external, to the subject beyond the mere occurrence of

¹⁶ For one who does, see Peacocke 2009.

perceptual experience are required for one to correctly grasp an instance as of the observable property it is. Equally, in the case of self-knowledge, I can err in thinking that I am thinking a thought with a certain content, since further conditions having to do with conscious attention and absence of interference from perceptual and other, standing, states are required for me to correctly grasp my first-order state as of the contentful property it is. In perceptual experience, I can misperceive an instance of red as of green (thereby misperceiving redness as greenness), and I can err when thinking that I am wondering whether Cologne is on the Rhine by grasping my actual first-order wondering whether Cologne is on the Seine *as* a wondering whether Cologne is on the Rhine. How and why this is possible is addressed in Section 4 below.¹⁷

Constitutivists like Shoemaker and Burge need not (and Burge in particular does not) deny the epistemic immediacy of the relation between a subject's thoughts about her thoughts and the contents of those thoughts when thinking and thinking about them. But they do both deny that this immediacy is anything like the kind of immediacy that attaches to cases of perception. My point is that we can all agree that the distinctive

¹⁷ As the claim suggests, there is a third way in which perception and self-knowledge in the core cases cited here are relevantly alike. Elsewhere (Macdonald 1999) I have discussed Shoemaker's (1994) arguments against the Object Perception Model and the Broad Perception Model. Although I cannot repeat it here, of particular relevance to the points made here is my argument for the claim that the Independence Thesis is not, as Shoemaker maintains, true in the case of perception but false in the case of self-knowledge. Many consider the *is/appears* distinction to be symptomatic of the independence of objects of perception from perceptual experience. If I am right, it is also symptomatic of the independence of one's first-order mental states from one's second-order ones in many cases of self-knowledge. For other reasons for thinking that Shoemaker's arguments do not succeed, see Peacocke's (2009) argument for the claim that the case of self-blindness does not establish that the Independence Thesis is false, which is compatible with an introspectionist, detectivist account of the kind I advance here.

way in which one knows about one's own mental states in cases of current conscious thinking is not the perceptual, experiential way without denying that both types of awareness are epistemically direct in structurally parallel ways.

The two features that make for direct epistemic access (i.e., being the favoured route for knowledge, and being as they appear to be), taken together, do not give one authoritative self-knowledge, but they do provide an epistemic basis for an introspective account of such knowledge. However, if subjects are the only ones to whom their contentful intentional properties *can* appear in the epistemically basic and direct way, then it follows that in these cases, barring cognitive failures, subjects have authoritative knowledge of contents of their own intentional states. Others too can have epistemic access to one's own states, but they grasp those states in the observational way. This asymmetry is what explains (but does not ground) the asymmetry between self-knowledge and knowledge of others that is symptomatic of first-person authority. And when I know my states in this observation-based way, my knowledge is not authoritative.

Section 3: Second-Order Thinking and the Temporal Issue

Section 1 argued there are different kinds of cases of conscious thinking about one's current, conscious thoughts where the *cogito*-like same-order interpretation does not do justice to the phenomenology of second-order thinking, as set out by the constraint (PC). Section 2 addressed part (a) of that constraint — that subjects' conscious thoughts about their current, conscious intentional states typically present those states to their subjects in an epistemically direct and immediate way — and used this as a basis for a characterisation of the special epistemic security that is characteristic of certain self-knowledge — in what that special security consists — that is

consistent with detectivism. This is not an explanation of *why* subjects are authoritative about their own conscious states. It is an explanation that is consistent with detectivism of what it is to have such authority, given that we do at least sometimes have epistemically direct and immediate access to our own conscious states. It remains to address part (b) — that subjects’ conscious thoughts about their current, conscious intentional states typically present those states *as* distinct from and present antecedently to thoughts about them in a way that attracts and engages their attention. Doing so will require addressing two issues: a temporal one, and a semantic one.

This section addresses the temporal issue. How is it possible to meet part (b) of the phenomenological constraint while also maintaining that, in cases of conscious thinking about one’s own current thoughts, one is *both* thinking and thinking *about* a given thought at the same time? Doesn’t thinking and thinking about a thought at the same time require that the first-order thought is a constituent of the second-order one (and so a *cogito*-like state)? If not, doesn’t thinking about a thought at a given time require that the thought thought about has already occurred at an earlier time, distinct from the time at which one is thinking about it (and so not at the same time)?¹⁸

This depends on how one understands the phrase ‘the same time’. Section 2 noted that situations involving occurrent propositional memory impressions and thoughts about one’s current musings, conjectures, wonderings, and the like are ones in which it is plausible to say that a first-order thought presents itself to a subject as being

¹⁸ This objection was famously pressed by Auguste Comte, who claimed,

But as for observing in the same way *intellectual* phenomena at the time of their actual presence, that is a manifest impossibility. The thinker cannot divide himself into two, of whom one reasons whilst the other observes him reason. (1830, using the translation of James 1890/1981:188)

James agrees that consciousness cannot be split, not because it is inconceivable but because it is empirically unlikely, and gives a different account of self-knowledge in terms of ‘retrospection’. See also Ten Elshof (2005).

present antecedently to and independently of her awareness of it and as attracting her attention in reflection. This indicates that the process of reflection *takes* time. In general events, being changes, take time; and mental events are no different than others in this respect. Given this, reflective thinking about one's current thoughts in the cases of interest here seems to involve what is commonly referred to as the specious present. According to one account of this, subjects experience events that are past, but so immediately so that they are experienced *as* present.¹⁹ According to another, subjects immediately experience change, or succession, during a period of time that is an interval (Dainton 2000). A familiar example often given of this is that of experiencing a melody, where the ability to have the experience seems to require the ability to experience a series of notes, occurring in succession over a period of time, as an interval which is present rather than as a durationless instant. Although the first note of the melody has passed by the time it is experienced *as* part of the melody, it, along with the other notes, are experienced *as* present. It is an experience of succession, and not merely a succession of experiences.

¹⁹ The concept of the specious present is due to the work of the psychologist E.R. Clay (quoted by James 1890) but its best known characterisation is due to William James (1890), who said "The prototype of all times is the specious present, the short duration of which we are immediately and incessantly sensible". According to Le Poidevin, the specious present is "...a duration which is perceived both as present *and* as temporally extended. This present of experience is 'specious' in that, unlike the objective present (if there is such a thing...), it is an interval and not a durationless instant. The real or objective present must be durationless for, as Augustine argued, in an interval of any duration, there are earlier and later parts. So if any part of that interval is present, there will be another part that is past or future" (Le Poidevin 2009). If knowledge of one's own current conscious states is understood as taking place during an interval that is the specious present (more accurately, during an interval that is a series of overlapping specious presents, since, as Dainton (2004) (citing Pöppel 1985) notes, the shortest interval during which human subjects can detect temporally ordered stimuli appears to be at most approximately a second), then one's reflective state is reflecting on a state which is, during this interval, past in 'real' terms. But what matters for the account is that it is experienced *as* present, and, as Le Poidevin points out, most states that are experienced as such are in the immediate past, and are so for good reasons (see Butterfield 1984). Perception of ordinary objects, such as trees, seems to involve the specious present, since the fact that both light and sound, as well as information travelling from receptors to the brain are transmitted at a finite speed seems to have the consequence that we can only perceive what is past, albeit the immediate past. What philosophers have called 'the specious present' seems to involve what neuropsychologists call 'working memory' (cf. Larner 2008). It is generally acknowledged that the concept of working memory is a descendent of that of the specious present and used by contemporary psychologists to indicate a dynamic temporary information storage system whose contents are available 'online' for current complex cognitive processing tasks.

The specious present is said to be specious because it is experienced, not as durationless, but as an interval. The process involved in reflective thinking about one's musings, speculations, wonderings and so on, are examples of this kind. What exactly is the relation between the first-order thought, the thought, '*p*', and the second-order thought, 'I am currently thinking that *p*', in this process? One suggestion stems from the analogy referred to above between this process and the process of experiencing a melody. Although the first note of the melody has passed, in 'real' time, by the time I experience it *as* part of the melody, we don't want to say that, at any given time during the interval in which I am having that experience, it is strictly speaking false that I am experiencing the melody because I am experiencing only part of it. In order to accommodate this intuition, we need to suppose that what it is for an experience to be a 'current' experience of a melody is for it to take up a period of time, however short, that includes the exact, or 'real' present, *as well as* the 'real' past (as suggested by the characterisation of the specious present as an immediate experience of change or succession during an interval of time). Analogously, it seems, in the case of reflection, we need to suppose that what it is for a second-order thought to be 'currently' present is for it to take up some time, however short, that includes the exact, or 'real' present – as well as the time at which the first-order thought, '*p*', occurs.

Further, just as the experience of the melody involves immediately experiencing the succession of notes constituting the melody during an interval of time, that experience being a 'current' one, and not just having a succession of experiences of those notes during that interval, so too thinking about one's own first-order thoughts involves being immediately conscious of those thoughts during an interval of time, that second-order process being a current one, and not just having a succession of thoughts during that interval, the first-order one, and the later, second-order one. The analogue

of the ‘experience’ of succession, in the thought case, is plausibly captured by the phenomenology of ‘aboutness’ that relates the later, second-order thought to the earlier, first-order one, in a direct and immediate way.

Section 4: Explaining Aboutness

This raises the second issue mentioned at the outset of section 3: how to capture the relevant notion of ‘aboutness’, or reference to the first-order content. Since it plays such a critical role in a robustly two-order view about the relation between first- and second-order thoughts in the kinds of cases under consideration here, how is it to be explained? One suggestion is that it is secured by attention; that one secures reference to the first-order thought just by attending to it (Hofmann 2009). However, while attention does seem to be an integral part of what is involved in conscious reflection, the notion seems too closely linked to that of reflection to illuminate it. For one thing, ‘attention’ is a success term, and a general account of the epistemic achievement involved in authoritative self-knowledge in the kinds of cases of concern here must allow for attempts that fail. Failed attempts call out for explanation, and notions such as those of inattention or insufficient attention are simply placeholders for an account of what it is that makes for success when success is the result, and what is missing when attempts to succeed fail. For another, there ought to be some connection between what makes for success and the semantics of the situation. The relevant notion of ‘aboutness’ that is in play when one speaks of one’s second-order thoughts being about one’s ontologically distinct first-order ones is presumably a semantic notion. But the notion of attention does not seem to be a semantic one.

A much more promising suggestion is that ‘aboutness’ or reference to the first-order content is secured in a way analogous to that suggested by Davidson’s

(1969/2001) paratactic analysis of sentences involving indirect discourse. Suppose that Sam sincerely asserts that Galileo said that the earth moves. According to the paratactic analysis, the sentence attributing this saying to Galileo has the form of an expression referring to a speaker (in this case, Galileo), a two-place predicate ‘said’ and a demonstrative referring to an utterance whose content is demonstrated and then produced by Sam but which is not semantically part of that sentence (in this case an utterance of ‘The earth moves’). If he is successful, Sam manages to truly attribute an utterance with a particular content to Galileo by securing reference to an utterance of his own with the same content (which may or may not be captured by a use of the same sentence). He manages to truly attribute an utterance to Galileo and *by this means* attribute a content to it. And he does this by referring to Galileo and using the two-place predicate ‘said’.

Turn now to the case of thoughts and self-attributions of them. Suppose that I first think, *salt is white*, and then think, *I am currently thinking that salt is white*. I think that I am currently thinking that salt is white. In thinking the second-order thought, I attribute a thought to myself by referring to myself, attributing to myself an attitude (thinking) rather than an utterance, and demonstrating, then thinking, a propositional content which is not semantically part of that thought, the content *salt is white*. This second-order thought is true if and only if the content demonstrated in that second-order thought, namely, *salt is white*, and the content of my first-order thought – *salt is white* – are the same. If successful, I manage to truly self-attribute a thought with a particular content by securing reference to the content of one of my own current (i.e., speciously present) thoughts that is distinct from but has the same content as that of the thought attributed. I truly attribute the first-order content by securing reference to the second-order content; but I only manage successfully to attribute a first-order

thinking of that content by both truly attributing to myself a thinking (which I do by thinking ‘I am thinking’ and by demonstrating a content which I then think and which has the same content as the thought attributed).²⁰

The present suggestion is a straightforward extension of the Davidsonian analysis of indirect discourse. This is not surprising, since both in indirect discourse and in the cases of self-knowledge of concern here there are two states whose contents need to bear a certain semantic relation to one another. The question might be raised, why not simply take the second-order thought to demonstrate the content of the first-order one rather than to take it to demonstrate its *own* content and by this means to secure reference to the content of the first-order thought? The reason is that there seems to be an aspect of the content of the second-order content, and not just the thinking of it, that is present when thinking about one’s first-order thought, presenting the content *as* the content of the first-order thought, and so as a candidate for reflection. I do not just demonstrate the content that is the same as the first-order thought content: I represent that content as a content that I am thinking (by first thinking ‘I am thinking’). Demonstrating the content of a first-order thought does not capture this aspect of second-order thinking. Further, and relatedly, simply demonstrating the content of another thought is not thereby to think *about* it, any more than demonstrating the content of another’s saying is to say something about it. In indirect discourse, I attribute a saying to another, not directly by demonstrating it, but indirectly, by

²⁰ See LePore and Loewer (1989) for an extension of the paratactic analysis to sentences ascribing mental states such as beliefs. One referee has asked for some further explanation of what is involved in demonstration of lower level content. It is difficult to say briefly here what more might be involved than is often claimed by those who say that there are mental demonstratives, conceptual analogues of linguistic demonstratives (see, e.g., Levine 2010). However, Jesse Prinz (2007) has argued persuasively for the view that there is good evidence for thinking that there are such demonstratives, “mental pointers”, and that they are required in an explanation of how subjects can think about their own phenomenal mental states. Although he distances himself from the view that these ‘pointers’ are conceptual items, preferring a view that takes them to be mechanisms that control top-down attentional focus, I favour a view that takes such items to be conceptual, taking attentional focus to be a further necessary but not itself a sufficient condition for mental demonstration (as I indicate in the text).

producing it myself; I *say* what another has said. When I think about my own thoughts, I attribute a thought to myself, not directly by demonstrating it, but indirectly, by *thinking* what I antecedently think.

This makes the view robustly second-order, but not second-order in the traditional way. A clear example of a position that involves genuinely second-order thought is a Fregean one, which utilizes a hierarchy of senses in order to make a clear use/mention distinction. A genuinely second-order thought understood along Fregean lines would merely represent a first-order content by mentioning it, rather than involve a content of the same type by using it. In contrast, the position articulated here involves a Davidsonian refinement of the use/mention distinction rather than introducing anything like a Fregean hierarchy.²¹

As a result, the position advanced here is very much like the Burgean same-order view, but with an important difference: this strategy replaces the reviewed first-order thought in Burge's theory with another that, under optimal circumstances, will have the same content as that reviewed thought. The replacement allows for ontological independence, fallibility, and distanced reflection on the thought reviewed.

²¹ It would seem to be a consequence of this view that when we think we are thinking a thought with content *p*, we replicate the content *p* in the second-order thought by both mentioning and using it, since this is the way in which we represent the contents of our first-order thoughts. But do we really notice two thoughts with the relevant contents? In the situation envisaged in the *Aida* example, am I really thinking 'I want to see *Aida*' twice? If so, is this compatible with the phenomenological constraint (PC) set out in section 1? Yes, I really am thinking 'I want to see *Aida*' twice and notice two thoughts with the relevant contents. And yes, this is compatible with (PC). (PC) tells us that subjects' conscious thoughts about their current, conscious intentional states typically present those states to their subjects as states that they are thinking *about* that are independent of and present antecedently to thoughts about them and as attracting and engaging their conscious attention. The phenomenon of thinking and thinking about is something that is part of the phenomenology of reflective thinking about one's first-order thoughts, and this means both that I notice that the content *p* (I want to see *Aida*) is present antecedently to my thought about it and also I am aware that my reflective thought is a thought *about* that content, i.e., that it presents the content *p* as the content of my first-order state. According to the paratactic account, the way in which I think about *p* is by representing it in my reflective thought, and I do this by both mentioning and using *p* itself. I cannot think the reflective thought without noticing its content, and that thought cannot present the first-order content one I am thinking *about* without presenting *p* as independent of and present antecedently to my reflective thought (with its own content). So the phenomenology of thinking and thinking about is consistent with noticing two contents, the content thought (in the reflective thinking) and the content thought about (in the thought reflected upon).

The Davidsonian model of how reference is secured in introspection not only predicts that there can be mismatches in content, but also, it seems, ones that we do not detect introspectively. Suppose, for example, that I hear John say that Joe likes Ella, but then later say that he said that Joe likes Stella. On the paratactic analysis, my saying is false because our sayings mismatch in content. Applied to the case of self-knowledge, the model predicts cases where, for example, I think I am wondering whether Cologne is on the Rhine but I am actually wondering whether Cologne is on the Seine, and my self-ascription is false because I fail to think a thought with the same content as the content with my first-order thought. But if I do, it seems that I do not detect this mismatch introspectively. But how can there be such undetected mismatches in content? Doesn't this show that the Davidsonian model of how reference is secured in introspection is false?

The challenge here is not simply to the view that there can be mismatches in content in introspective awareness. This is something that detectivism must permit given the contingent causal connection between ontologically distinct first- and second-order thoughts, and many detectivists (e.g., Armstrong, Lycan) openly acknowledge that such errors are possible. The challenge is specifically to the thought that such mismatches can go undetected introspectively *at the time of introspecting*. Now, it is tempting to think that one can respond simply by pointing out that undetected errors due to mismatches in content are rife throughout the cognitive domain. Consider the case of hearing. It is true that when I hear John and say that Joe likes Ella but then later say that he said that Joe likes Stella, there is a difference in the times of the two sayings. John's saying occurs at a time before the time of my attribution to him of a saying. However, when I err in reporting what John said, I do not detect the mismatch in content between what I hear and what I report. Nor does

my audience. The detection of the mismatch by an audience requires both knowing the content of what John said and also hearing my attribution of a saying to John. That audience is in a position to correct the mismatch in content, but it can do so only *after* hearing my attribution, and only given the knowledge of the content of what John said. The thought that one might detect a mismatch in content in the case described here, then, is itself problematic; my audience must rely on memory to identify the mismatch, since John's saying occurs at a time prior to the time that I make my attribution. And if I detect a mismatch at some later time, I too must rely on memory. So, it seems one cannot detect a mismatch in the case of indirect discourse. Why should we expect introspective mismatches in content to be detected when such errors in other areas of our cognitive lives are not?

The reason is that there appears to be a crucial difference between the case of introspection and that of hearing. When I hear John say that Joe likes Ella and say that he said that Joe likes Stella, the mismatch in content occurs over a period of time. Memory is involved and memory can be defective. Equally, and comparably, I can easily falsely think that I *was* wondering whether Cologne is on the Rhine, since here too there is a time difference between my wondering whether Cologne is on the Seine and my thinking that I was wondering whether Cologne is on the Rhine; memory is involved, and my memory can be defective. The difference, in the case of my thinking that I am wondering whether Cologne is on the Rhine, is that my self-ascription is in the *present* tense, and it is this feature that makes it seem impossible that I should err and not detect it. The difficulty in describing cases where there are undetected introspective errors fosters the view that there is a much more intimate connection between first- and second-order thoughts in cases of current conscious thinking than the detectivist view can acknowledge.

However, the cases are not as different as the objection makes them out to be. Everything depends on how the details of the detectivist position are developed. The one developed here proposes a solution to the temporal problem – of how one can both think and think about a given content at the same time —by appeal to the doctrine of the specious present. This is a unique and key feature of the account. The temporal problem is resolved by recognizing that introspecting is a process that takes time – that the conscious present is not a durationless moment — and the proposal is that one’s thoughts about one’s current conscious thoughts are to be viewed as taking place during an interval of time, however short, that is a series of overlapping specious presents.

As noted in Section 3, James used the concept of the specious present to refer to what contemporary theorists call working memory. Although there is some controversy about the various features and functions of working memory, common to the differing characterisations is that it is a dynamic information storage system that stores a limited amount of information for a brief period of time that is quickly accessible for higher cognitive processing and is subject to frequent updating.²² Working memory is sometimes described as a system of ‘online’ information storage whose capacity and functioning varies from individual to individual. It is responsible for supporting minute-to-minute monitoring and processing in a highly varied range of cognitive tasks, some of which include thinking, problem solving, and reasoning.

There are issues of debate amongst researchers concerning the scope and functions of working memory (cf. Baddeley 1986, 1992), the nature its limitations,

²² See Jonides et al (1996). So ‘working memory’ refers to a system that both temporarily stores *and* processes information. This, amongst other features, distinguishes it from short-term memory (Dehn 2011, Hutton and Towse 2001), which is taken to be a passive storage system. There is a good deal of dispute about how long exactly the period of time is that the information in working memory is available for cognitive processing; views range from 1-2 seconds, to roughly 15 seconds, to 30 seconds.

exactly what role it plays in complex cognition, and, importantly for present purposes, its relationship to attention and consciousness.²³ What is not generally disputed is that much of working memory is conscious, even if not all of it is. Nor is it disputed that there is an intimate relation between working memory and attention, and that working memory load can affect working memory task performance (Broadway and Engle 2011).

The specious present for James referred mainly to the consciousness or awareness subjects have of information to which they consciously attend. Irrespective of how the above disputes are resolved, then, the account of self-knowledge developed here concerns that part of working memory that is conscious and attended to, and thoughts about one's current conscious thinkings requires and engages that part of working memory.

The relevance of the foregoing to the objection under consideration is this. Despite the fact that the Davidsonian model of indirect discourse concerns attributions in the past tense whereas self-attributions in the relevant cases here are in the present tense, there is a structural parallel between the two types of case. Specifically, just as attributions of content in indirect discourse in the past tense rely on memory, self-attributions of content in the present tense rely on working memory. Further, and importantly, just as memory can be faulty, so too can working memory. In neglect, for example, there is a deficit of attention, and this can affect working memory. The

²³ For examples of the view that the contents of working memory are the contents of consciousness, see Moscovitch & Umiltà (1990) and Atkinson and Shiffrin (1971). For examples of the view that the contents of working memory 'outstrip' the contents of consciousness, see Baars' (1997) Global Workspace theory, according to which only those items in working memory that are the focus of attention can be consciously experienced, and Cowan (1988). This debate is intimately connected with the debate concerning whether there is consciousness outside attention. (Cf. Prinz (2000, 2010, 2012) for the view that attention is both necessary and sufficient for consciousness, Chalmers (1996, 2010) and Block (1995) for the view that consciousness outstrips attention.)

functioning of working memory can also be affected by forgetting (Baddeley 1986, 2007).

Crucially, the reliance of present-tense self-attributions on working memory provides a ready explanation for there being mismatches in content that are not detected introspectively at the time of introspecting. Just as there is a time difference between my hearing John's utterance that Joe likes Ella and my saying that John said that Joe likes Stella, there is a time difference, however short, between my wondering whether Cologne is on the Seine and my thinking that I am wondering whether Cologne is on the Rhine. This being so, the mismatch of content can go undetected introspectively in my own case just as the mismatch of content can go undetected in the case of my attribution to John of a saying with a content different from the one I heard him say, despite the fact that the self-attribution is in the present tense. The fact that self-attributions based on introspection are made in the present tense *masks* this time difference and can make it appear that subjects are immune from error in the kinds of cases of current conscious thinking about one's thoughts that are of concern here. But this appearance is misleading; subjects are no more immune from error in such cases than they are in cases of indirect discourse. In neither case are mismatches in content detected, and in both cases one must rely on memory (and not merely on working memory) and other background and current knowledge to identify such mismatches.

There is an analogue of introspectively undetected mismatches in content in perception, given that perception itself seems to involve the specious present or its descendant, working memory (cf. n.19). In perception, mismatches between perceptual content and what such content represents when misperceiving go undetected at the time that the misperception occurs, and require reliance on memory to detect.

Similarly, I suggest, mismatches in content in self-knowledge that are undetected at the time of introspecting require reliance on memory to detect. In both cases, the source of the possibility of such undetected mismatches is the same – the contingent, but reliable, causal connection between ontologically distinct states. Introspection is fallible, despite its giving direct and immediate access, *when it does give such access*, to the contents of our first-order states. What the paratactic analysis tells us is that subjects can have direct and immediate access to the contents of their first-order states via thinking the contents of our second-order states, parts of which represent those first-order contents by replicating them. When all goes well, the first-order contents are transparent to second-order reflection because there is sameness of content. Given the reliability of the causal connection between reviewing thoughts and thoughts reviewed, we have every reason to think that this is the default position. The reliability of the connection gives no more reason to expect mismatches in content to occur with any more regularity than they do in the case of perception. We are right to trust the deliverances of our perceptual mechanisms, even if they can sometimes mislead us as to what we are perceiving. Equally, we are right to trust the deliverances of introspection, even if they can sometimes mislead us as to what we are currently thinking.

Section 5: Putting Introspectionism and Detectivism Together

In the light of the foregoing, my proposal about the nature of the states related as reflecting and reflected upon, and their relation to one another, in the kinds of cases that form the target class is as follows:

When I think about my first-order intentional state in an epistemically direct or immediate way:

1. I do so during two time intervals that are overlapping specious presents.
2. The second-order intentional state fallibly attributes a content to one of my own first-order states by demonstrating, then thinking, it.
3. The first-order state – the current, conscious intentional state about which I am thinking and on which I am reflecting – is ontologically distinct from and causally related to this second-order state.
4. This being so, although I am better placed than others to know the content of my first-order state and so to have authoritative knowledge of it, that knowledge is neither incorrigible nor infallible.

In order for knowledge of one's own thoughts to be corrigible or fallible, the thought content known must be the first-order one - the thought content that is causally related to and ontologically distinct from the second-order thought attributing it. And it is. The extension of the paratactic analysis helps to show how it is possible for a subject to truly self-attribute a thought by thinking a second-order thought about a first-order one that is causally related to but ontologically distinct from it, just as the paratactic analysis itself helps to show how, in indirect discourse, a speaker can truly attribute a saying to another.

How is authority secured? When a subject is thinking about her own conscious intentional state, she is the only one who thinks a content of a type that can be known in an epistemically basic and direct way and in this instance *is* known in that way. As noted in section 4, when successful, the content of the reflective thought is, and presents itself *as* having (by representing that content as a content that I am thinking) a constituent that is partly of the same type as that which forms the content of the thought reflected upon. The content of the thought reflected upon is one to which the subject has direct epistemic access (i.e., it is epistemically basic and as it appears to the

subject to be). So the intellectual moves involved in recapturing an earlier train of thought no more involve a process of inference than do those involved in thinking *cogito*-like thoughts themselves. One either grasps the contents of the earlier states, or one does not. If one does, it cannot be by thinking thoughts with different contents. So if one does, one grasps the contents of the earlier states with as much epistemic immediacy as one does in thinking the contents of the reflective states.

The overall position is, as I have claimed, an introspectionist, detectivist one. What makes it introspectionist is that the second-order, reflective states, are *about*, and have the same contents as, one's own first-order psychological states, of which one is aware in a direct and immediate way. Its affiliations with detectivism are reflected in the recognition of a causal relation between first-order and second-order states that are ontologically independent of one another. This allows for both fallibility and distanced reflection on one's own intentional states.²⁴

University of Manchester
Manchester M13 9PL
United Kingdom

²⁴ I would like to thank Tim Bayne, Eve Garrard, Kathrin Glüer-Pagin, Ben Jarvis, Lawrence Lombard, Graham Macdonald, Asa Wikforss, Daniel Whiting, and audiences at the Universities of Glasgow, Hertfordshire, Southampton, Dublin, Canterbury, New Zealand, and anonymous referees for this journal, for very helpful comments.

References

- Alston, W. 1971. 'Varieties of Privileged Access'. *American Philosophical Quarterly*, 8: 223-41.
- Armstrong, D.M. 1981. *The Nature of Mind and Other Essays*. Ithaca, New York: Cornell University Press.
- Atkinson, R. C. and Shiffrin, R. M. 1971. 'The Control of Short-Term Memory'. *Scientific American* 225: 82-90.
- Audi, R. 2004. 'Perception and Consciousness'. In I. Niiniluoto, M. Sintonen, and J. Wolenski (eds.), *Handbook of Epistemology*. Dordrecht: Kluwer Academic Publishers, pp. 57-108.
- Baars, B. J. 1997. *In the Theater of Consciousness: The Workspace of the Mind*. New York: Oxford University Press.
- Baddeley, A. D. 1986. *Working memory*. New York: Oxford University Press.
- Baddeley, A. D. 1992. 'Consciousness and Working Memory'. *Consciousness and Cognition* 1: 3-6.
- Baddeley, A. D. 2007 *Working Memory, Thought, and Action*. New York: Oxford University Press.
- Bar-On, D. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Clarendon Press.
- Bilgrami, A. 2006. *Self-Knowledge and Resentment*. Cambridge, MA: Harvard University Press.
- Bilgrami, A. 2010. 'Replies to Tom Baldwin and Calvin Normore'. *Philosophy and Phenomenological Research* 81: 783-808.
- Block, N. 2003. 'Mental Paint'. In M. Hahn and B. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge, Mass.: MIT Press, pp. 165-200.
- Block, N. 1995. 'On a Confusion About a Function of Consciousness'. *Brain and Behavioral Sciences* 18: 227-247.
- Block, N. 2007. 'Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience'. *Behavioral and Brain Sciences* 30: 481-500. Commentaries and Replies: 500-548.
- Boghossian, P. 1989. 'Content and Self-Knowledge'. *Philosophical Topics* 17: 5-26.
- Boghossian, P. and Velleman, J. D. 1991. 'Physicalist Theories of Color'. *Philosophical Review* 100: 67-106.

- Boyle, M. 2009. 'Two Kinds of Self-Knowledge'. *Philosophy and Phenomenological Research* 78: 133-164.
- Broadway, J. and Engle, R. 2011. 'Lapsed Attention to Elapsed Time? Individual Differences in Working Memory Capacity and Temporal Reproduction'. *Acta Psychologica* 137: 115-126.
- Burge, T. 1996. 'Our Entitlement to Self-Knowledge'. *Proceedings of the Aristotelian Society* 96: 91-116.
- Burge, T. 1998. 'Reason and the First Person'. In C. Wright, B. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 243-70.
- Burge, T. 1999. 'A Century of Deflation and a Moment about Self-Knowledge'. *Proceedings and Addresses of the American Philosophical Association* 73(2): 25-46.
- Butterfield, J. 1984. 'Seeing the Present'. *Mind* 93: 161-76. (Reprinted with corrections in R. Le Poidevin (ed.), *Questions of Time and Tense*. Oxford: Clarendon Press, 1998, 61-75.
- Campbell, J. 1994. 'A Simple View of Color'. In J. Haldane and C. Wright (eds.) *Reality, Representation and Projection*. Oxford: Clarendon Press, pp. 257-69.
- Chalmers, D. 1996. *The Conscious Mind*. New York: Oxford University Press.
- Chalmers, D. 2010. *The Character of Consciousness*. New York: Oxford University Press.
- Coliva, A. 2009. 'Self-Knowledge and Commitments'. *Synthese* 171: 365-375.
- Comte, A. 1830. *Cours de philosophie positive*, volume 1. Paris: Bacheleier, Libraire pour les Mathématiques.
- Cowan, N. 1988. 'Evolving Conceptions of Memory Storage, Selective Attention, and Their Mutual Constraints Within the Human Information Processing System'. *Psychological Bulletin* 104: 163-191.
- Dainton, B. 2000. *Stream of Consciousness: Unity and Continuity in Conscious Experience*. London: Routledge.
- Dainton, B. 2004. 'The Self and the Phenomenal'. *Ratio* 14(4): 365-89.
- Davidson, D. 1969/2001. 'On Saying That'. Reprinted in his *Inquiries Into Truth and Interpretation: Philosophical Essays*, Vol. 2. Oxford: Clarendon Press, 93-108.
- Dehn, M. 2011. *Working Memory and Academic Learning: Assessment and Intervention*. New Jersey: John Wiley & Sons.
- Dretske, F. 1969. *Seeing and Knowing*. London: Routledge and Kegan Paul.

- Dretske, F. 1978. 'The Role of the Percept in Visual Cognition'. In W. Savage (ed.), *Perception and Cognition: Issues in the Foundations of Psychology*. Minneapolis: University of Minnesota Press, 107-27.
- Dretske, F. 1979. 'Simple Seeing'. In D. Gustafson and B. Tapscott (eds.), *Body, Mind, and Method*. Cambridge, Mass.: MIT Press, 1-16.
- Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Evans, G. 1982. *The Varieties of Reference*. Ed. by J. McDowell. Oxford: Oxford University Press.
- Finkelstein, D. 2008. *Expression and the Inner*. Cambridge, Mass.: Harvard University Press.
- Gertler, B. 2011. 'Self-Knowledge and the Transparency of Belief'. In A. Hatzimoysis (ed.), *Self-Knowledge*. Oxford University Press, pp. 125-145.
- Heil, J. 1983. *Perception and Cognition*. California: University of California Press.
- Hofmann, F. 2009. 'Introspective Self-Knowledge of Experience and Evidence'. *Erkenntnis* 71: 19-34.
- Horgan, T. and Tienson, J. 2002. 'The Intentionality of Phenomenology and the Phenomenology of Intentionality'. In D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press, 520-33.
- Horgan, T., Tienson, J., and Graham, G. 2004. 'Phenomenal Intentionality and the Brain in a Vat'. In R. Schantz (ed.), *The Externalist Challenge*. Berlin: Walter de Gruyter, 297-317.
- Hutton, U. and Towse, J. 2001. 'Short-Term Memory as Indices of Children's Cognitive Skills'. In S. Gathercole (ed.), *Short-Term and Working Memory*. Sussex: Psychology Press, pp. 383-394.
- James, W. 1890/1981. *The Principles of Psychology*. Cambridge, Mass.: Harvard University Press.
- Johnston, M. 1992. 'How to Speak of the Colors'. *Philosophical Studies* 68: 221-263.
- Jonides, J. et al. 1996. 'Verbal and Spatial Working Memory in Humans'. *The Psychology of Learning and Motivation* 35: 43-88.
- Kind, A. 2003. 'What's So Transparent About Transparency?' *Philosophical Studies*: 225-244.
- Larner, A. J. 2008. *Neuropsychological Neurology: The Neurocognitive Impairments of Neurological Disorders*. Cambridge: Cambridge University Press.

- Le Poidevin, R. 2009. 'The Experience and Perception of Time'. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*. URL= <http://plato.stanford.edu/archives/win2009/entries/time-experience/>.
- LePore, E., and Loewer, B. 1989. 'You Can Say *That* Again'. In *Midwest Studies in Philosophy* 14: 338-56.
- Levine, J. 2010. 'Demonstrative Thought'. *Mind and Language* 25(2): 169-95.
- Lycan, W. 1996. *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Macdonald, C. 1998. 'Externalism and Authoritative Self-Knowledge'. In C. Wright, B. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 123-154.
- Macdonald, C. 1999. 'Shoemaker on Self-Knowledge and Inner Sense', *Philosophy and Phenomenological Research* 59: 711-738.
- Macdonald, C. 2007. 'Introspection and Authoritative Self-Knowledge'. *Erkenntnis* 67: 355-372.
- McNaughton, D. 1988. *Moral Vision*. Oxford: Basil Blackwell.
- Moscovitch, M. and Umiltà, C. 1990. Modularity and neuropsychology: Modules and central processes in attention and memory. In M. F. Schwartz (ed.), *Modular Deficits in Alzheimer-Type Dementia*. Cambridge, MA: MIT Press, pp. 1-59.
- Nudds, M. and O'Callaghan, C. 2009. 'Introduction'. In M. Nudds and C. O'Callaghan (eds.), *Sounds and Perception*. Oxford: Oxford University Press, 2009, pp. 1-24.
- O'Callaghan, C. 2007. *Sounds*. Oxford: Oxford University Press.
- Parent, T. 2007. 'Infallibilism About Self-Knowledge'. *Philosophical Studies* 133: 411-424.
- Peacocke, C. 1998. 'Conscious Attention, Attitudes, and Self-Knowledge'. In C. Wright, B. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 63-98.
- Peacocke, C. 1999: *Being Known*. Oxford: Oxford University Press.
- Peacocke, C. 2009. 'Mental Action and Self-Awareness (II): Epistemology'. In L. O'Brien and M. Soteriou (eds.), *Mental Action*. Oxford: Oxford University Press, pp. 192-214.
- Pöppel, E. 1985. *Mindworks: Time and Conscious Experience*. New York: Harcourt Brace Jovanovich.
- Prinz, J. 2007. 'Mental Pointing: Phenomenal Knowledge without Phenomenal Concepts'. *Journal of Consciousness Studies* 14: 184-211.

- Prinz, J. 2010. 'When is Perception Conscious?'. In B. Nanay (ed.), *Perceiving the World: New Essays on Perception*. New York: Oxford University Press, pp. 310-332.
- Prinz, J. 2012. In Press. *The Conscious Brain*. New York: Oxford University Press.
- Russell, B. *The Problems of Philosophy*. London: Oxford University Press, 1912.
- Shoemaker, S. 1994: 'Self-Knowledge and Inner Sense'. *Philosophy and Phenomenological Research* 54: 249-314.
- Shoemaker, S. 1996: 'On Knowing One's Own Mind'. In S. Shoemaker *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press): 25-49.
- Shoemaker, S. 2009. 'Self-Intimation and Second Order Belief'. *Erkenntnis* 71: 35-51.
- Ten Elshof, G. 2005. *Introspection Vindicated*.
- Williamson, T. 1995. 'Is Knowing a State of Mind?'. *Mind* 104: 533-565.
- Williamson, T. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Wright, C. 1998. 'Self-knowledge: The Wittgensteinian Legacy'. In C. Wright, B. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 15-46.
- Wright, C. 2001. 'On Making Up One's Mind: Wittgenstein on Intention'. In C. Wright, *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*. Cambridge, MA: Harvard University Press, 2001: 116-142.
- Zimmerman, A. 2008. 'Self-Knowledge: Rationalism vs. Empiricism'. *Philosophy Compass* 3/2: 325-352.