

The peculiarity and contingency of the introspection of belief¹

A peculiaridade e contingência da introspecção da crença

Leandro De Brasi²
Alberto Hurtado University

Abstract

The Causal Model of introspection has its more than fair share of critics and indeed in recent years the model has fallen out of favor in the philosophical world. In this paper, I defend the model and argue that it is an excellent candidate, given a realist commitment about the mental, to explain our peculiar, but contingent, introspective access to beliefs.

Keywords: causal model, introspection, belief, peculiarity of self-ascriptions, contingency of self-ascriptions.

Resumo

O modelo causal de introspecção tem recebido muitas críticas e, de fato, nos últimos anos, tem caído em desuso no mundo filosófico. Neste artigo defendo o modelo e sustento que ele é um excelente candidato, dado um compromisso realista para com o mental, para explicar o nosso acesso introspectivo peculiar, mas contingente, às crenças.

Palavras-chave: modelo causal, introspecção, crença, peculiaridade de autoatribuições, contingência de autoatribuições.

Self-ascriptions of beliefs are thought to be peculiar, in the sense that they are arrived at differently from other-ascriptions. In this paper I defend the often questioned Causal Model of introspection and argue that it is an excellent candidate, given a realist commitment about the mental, to explain our peculiar, but contingent, introspective access to beliefs.

¹ Thanks are due to participants of the Third PLM Conference, the UAH Mind, Language & Knowledge Reading Group and an anonymous referee. Research for this article has been partially supported by the Fondecyt project #11140279 (Chile).

² Alberto Hurtado University. Departamento de Filosofía, Alameda 1869, 3er piso, 8340576, Santiago, RM, Chile. E-mail: ldebrasi@uahurtado.cl

The paper proceeds as follows. After some preliminaries (§§ 1-2), I introduce the phenomenon to be explained: the peculiarity of the introspection of belief (§§ 3-4). Next I introduce the Causal Model and show how it captures such peculiarity (§§ 5-6). After defending the model against objections regarding the contingency to which it is committed (§§ 7-8), I offer some concluding remarks (§ 9).

Realism about the Mental

It can prove difficult to know where to start in philosophy but, as Stroud (2000, p. 3) says, “a major part of philosophy as I understand it is the attempt to find out where that is.” I take realism about the mental to be the starting point for the present investigation and this paper can be partly seen as an attempt to show that this is a fruitful starting point. No doubt this starting point puts me at odds with some competing approaches (as we shall see), but I take this anyway to be an independently plausible starting point. It is nevertheless conceivable that realism about the mental is false, although its epistemic security seems greater than that of its competitors. Anyway it is my starting point, so debate about its truth isn’t really an issue here.

So let me explain what I understand by realism and what I take its consequences for an account of introspection to be. Realism is a metaphysical doctrine about what exists. I take it that the two basic (and related) tenets of a realist thesis about any area of discourse are: the claim that the objects of such discourse exist objectively and the claim that they exist independently from the mental.³ This mind-independence might seem to rule out the very possibility of realism about the mental, since the existence of the mental is, in a trivial way, mind-dependent. However, the mind-independence condition is meant to be understood as the claim, concerning the mental, that the world which contains mental states does not depend on the creatures’ capacity to determine that such world does contain such states.

So let us examine further the two general claims. Realism about the mental does not maintain that mental states are in no way subjective, but that they are *also* in some sense objective. Such objectivity is secured by the existence of an independent third-person perspective, the perspective from which we ascribe mental states to others, on the basis of observable behaviour (Smith, 1998). So it is this public character of the mental (i.e., mental items can in principle be observed by persons other than their possessors) which guarantees its objectivity and opens a potential gap between reality and appearance. That is, a potential mismatch between, say, one’s beliefs about one’s mental states and others’ beliefs about those states.

This commonplace appearance/reality gap introduces the possibilities of error and ignorance. On the one hand, ignorance is possible because one could lack contact with reality. And on the other, error is possible because one could make imperfect contact with reality. So in order to avoid error and ignorance with regard to substantive propositions of the discourse, one has to make adequate contact with reality. Indeed, the mind is there to be discovered rather than invented by our inquiry into it (cf. Tanney, 2002). And since discovery is a matter of contingent success, there is no guarantee that one will discover such reality.

Moving onto the second tenet, the mental can exist independently of one’s thoughts, judgements, beliefs, etc., about it. And one might be in a mental state without having any thoughts, judgements, beliefs, etc., about it or while having the

³ There are different notions of realism (e.g., Dummett, 1978), so I do not expect this characterisation to be universally accepted.

wrong ones. But the extent to which we know about such reality is an epistemic issue that I won't address here. The paper is concerned with *how* self-ascription (of belief) is arrived at, not whether we have self-knowledge or not.⁴ However, I do take the possibility of self-error and self-ignorance seriously. It is important that our account of self-ascription accommodates them: we need to allow for both success and failure in self-ascription.

So, given realism about the mental, mental states and their self-ascriptions are ontologically independent: one's mental states exist independently of their self-ascriptions. In other words, mental states do not depend for their existence on one's capacity to self-ascribe them; hence the need to discover them even in our own case. This contingent discovery points to a non-deflationary view of introspection.⁵ That is, introspective awareness is a genuine cognitive achievement: self-ascriptions track their ontologically distinct self-ascribed states (Fricker, 1998).⁶ So I take it that an account of introspection also needs to allow for introspection to be taken as a genuine cognitive achievement.

It is clear then that, given a commitment to realism about the mental, an account of introspection that accommodates these constraints is preferable. But as one might expect, the above constraints aren't universally accepted. So, for example, Cartesian approaches that deny self-ignorance and self-error (e.g., Descartes, 1984) and constitutivist approaches that deny the ontological independence of the self-ascription and the self-ascribed state (e.g., Shoemaker, 1994d) are in tension with this commitment.⁷

The Nature of Belief

Although I have so far been talking of the mental generally, my main concern in this paper is with belief. So here I briefly introduce the mental phenomenon of belief.⁸ To that purpose, and following Richard Wollheim (1999), I sketch a broad (though not exhaustive) map of the mental (more specifically, of the phenomena that our folk-psychological concepts pick out). The mental can be divided into two major categories: the *occurrent, experiential* phenomena (such as perceptions, sensations and entertained thoughts) and the *dispositional, non-experiential* phenomena (such as beliefs, desires and intentions). The fundamental difference between these two

⁴ Although I am not concerned with scepticism about self-knowledge in this paper, notice that realism about the mental does rule out one particular kind of scepticism about our minds: namely, the scepticism promoted by eliminativism. Briefly, eliminativism is the thesis that denies that our mental discourse picks out any real psychological categories (and so urges that such discourse be eliminated, hence its name; see Churchland, 1981).

⁵ In deflationary approaches (cf. Wright, 1998), self-ascriptions are not made by tracking some independent reality. That is, there is no genuine cognitive achievement (Gertler, 2008).

⁶ Fernandez (2013, p. 28) takes cognitive achievement to be "the result of some cognitive operation I perform." This is something done at the personal level and which takes an effort on my part: it is not something that happens to me (2013, p. 28). I take this to be an implausibly strong understanding of cognitive achievement if self-ascriptions are meant to represent cognitive achievements given that most self-ascriptions do not satisfy such characterisation (consider the case of self-ascriptions regarding stored beliefs). This notion of cognitive achievement also has undesirable consequences outside self-ascription (consider the case of perceptual beliefs). And these worries become more significant if one holds, as many do, that knowledge involves a cognitive achievement (e.g., Greco, 2010; Sosa, 2007).

⁷ Another approach that (at least a brand of) realism about the mental is in tension with is the expressivist one. Roughly, the expressivist claims that self-ascriptions are natural self-expressions (such as "Ouch!", where I'm simply expressing a painful state) that lack propositional content and so denies that self-ascriptions (beliefs/judgements about beliefs) are truth-evaluable (e.g., Wittgenstein, 1953; cf. Bar-On, 2004). A realist (like Dummett), on the other hand, holds that the discourse in question is apt for truth or falsity, which is determined by some set of facts. The self-ascriptions function in a genuinely descriptive way: they (are intended to) pick out the things there are. So the expressivist approach isn't even an option, given this realism about the mental.

⁸ Anti-realist conceptions of belief, such as the instrumentalist one (e.g., Dennett, 1987), are ignored.

categories is that while the former set of phenomena is properly characterised as mental acts (or events) that possess phenomenology, the latter is not.

Beliefs are dispositional states that lack phenomenology. First, beliefs are enduring, as opposed to passing, phenomena. Beliefs are not things that happen to one; they are not occurrences. Second, beliefs are not consciously experienced. There is no such thing as an 'occurrent belief' or a 'conscious belief'. As Crane says:

[We] need to distinguish between *being conscious of what you believe*, and *consciously believing*. [...It] makes little sense to say 'I have been believing p all morning'; or 'I believed p for a few hours last night and then I was distracted by the television'. [... So,] although there is such thing as being conscious of one's belief, that does not mean that there is such a thing as consciously believing (2001, p. 105-108).

Beliefs, unlike the occurrent, experiential states in which they manifest themselves (say, a judgement or an episode of affirmation), do not possess phenomenology. So accounts of introspection that detect mental states by means of their phenomenological properties, such as Alvin Goldman's phenomenological view (1993), cannot be applied to belief.⁹

Moreover, beliefs are propositional attitudes. Propositional attitudes are mental phenomena (of either category) which are formulated as 'S As that p', where 'S' refers to the subject, 'A' expresses the general type of attitude (say, believes, judges, hopes), and 'p' expresses a proposition.¹⁰ One can then distinguish between mental *content*, given by p, and *relation*, given by A, and so speak of propositional attitudes as having a relational nature, since they are relations to contents.

The mental relation is the characteristic type of role (or function) that differentiates each attitude. These relations can be divided into two kinds given their direction of fit. The belief relation has a *mind-to-world* direction of fit (the mind is to fit the world), as opposed to a *world-to-mind* direction of fit (the world is to fit the mind—consider, e.g., desires). Beliefs are the kind of disposition whose role is to provide one with a picture of reality. Indeed, we value our beliefs because we take them to be true. So the belief relation involves a commitment to the truth of the proposition believed. Beliefs are not just truth-apt (given their propositional content), they aim at truth.

The Introspection of Beliefs

The self-ascription of beliefs can be achieved in two different ways. It can be achieved through *self-interpretation* and through *introspection*. By 'self-interpretation' I mean the process of reasoning based on evidence by which we sometimes self-ascribe all sorts of mental states. So I understand self-interpretation as a Rylean-like type of self-ascription (Ryle, 1949). I say Rylean-like because the evidence on which the inferences are based needn't be only behavioural. Our self-interpretative ascriptions can also follow from premises about our personality traits and other mental states.

⁹ Indeed, here I am only concerned with accounts of introspection that are aimed at dispositional states or accounts that can be extended to them. See also Klausen (2008, p. 446-447), who argues for the thesis that propositional attitudes are essentially phenomenal, but restricts it so as to leave beliefs out (cf. Churchland, 1992, p. 32).

¹⁰ There are many views about the nature of propositions but, to a considerable extent, the debate about the nature of propositions runs separately from the debate about the nature of belief and so in this paper I remain silent on the issue.

Now, Paul Boghossian puts forward an argument against models of self-knowledge that base *all* self-ascriptions on inferences from premises about behaviour. His Anti-Rylean argument (Boghossian, 2003 [1989], p. 67) goes as follows:

You think: Even lousy composers sometimes write great arias. And you know immediately on thinking it, that that is what you thought. What explanation can the Rylean offer of this? The difficulty is not merely that, contrary to appearance and the canons of epistemic practice, he has to construe the knowledge as inferential. The difficulty is that he has to construe it as involving inference from premises about behaviour that you could not possibly possess.

Many have echoed this criticism of the Rylean approach. Usually writers point out that, even when one is perfectly still and silent, one can still know one's own mental states. That is, although there is no behaviour to determine what the mental state is, we still know it. But, as seen, there is no reason for the sympathiser of the Rylean approach not to extend the evidential resources available for the ascriptions. Indeed, any serious account will likely do so. So, although Boghossian's argument doesn't rule out the possibility of there being *only* (inferential and evidential) self-interpretative access to our mental states, it does shift the onus onto those who favour such views. They need to find a source of evidence that can account for *all* self-ascription and that has proven rather difficult to do.¹¹

So I take it that self-ascription can also be achieved through some other means: introspection.¹² I employ the term 'introspection' in a broad way so as to pick out whichever mental process is involved in the non-self-interpretative self-ascription of mental states. I am in this paper primarily concerned with this form of self-ascription because it is claimed to be special, in the sense of being epistemically distinctive. The claim that our introspective self-ascriptions are special is normally understood in terms of *privileged access*. But there are many different senses of the term (e.g., Alston, 1971; Gallois, 1996; Gertler, 2003). I differentiate here between two main epistemic claims: on the one hand, the security of self-ascriptions and, on the other, the peculiarity of self-ascriptions (see also Wright, 1998; Fernandez, 2003; Gertler, 2010; Smithies and Stoljar, 2012). Roughly, the latter maintains that self-ascriptions can be known *differently* from the way we know the ascriptions to others; the former states that self-ascriptions can be known *better* than the ascriptions to others.¹³ Here I focus on the peculiarity thesis. So let us consider how self-ascription might be peculiar.

¹¹ For example, Gopnik (1993) suggests that evidence to be psychological experiences, but we can easily conceive of cases where no such experience is involved in the self-ascription (indeed, Boghossian thinks that a similar argument can be given for any source of evidence suggested). So Gopnik's additional source of evidence doesn't do the job and the prospects of there being one are gloomy. In fact, if you, like me, think that some self-ascriptions are groundlessly warranted—there are cases where it would be unreasonable to ask someone "How do you know that you believe that?" (see, e.g., Neta, 2011, p. 18-19), then Boghossian might just be right (cf. Cassam, 2014).

¹² Although self-interpretation does not seem able to accommodate all self-ascription, we cannot give it up on pain of having no chance to capture confabulation phenomena (see, e.g., Carruthers, 2011). Furthermore, phenomenologically speaking, it seems pretty clear that sometimes we ascribe mental states that way (see, e.g., Cassam, 2014).

¹³ Peculiarity and security are two independent claims. So, for example, Ryle (1949) denies the peculiarity claim (our access to our own minds is the same in kind as our access to other minds) while accepting the security claim (we have more access to the relevant evidence in our case than in the case of others); so he rejects only one of the two claims associated with privileged access. The reverse case is also possible: we might have access to our own mind in a peculiar way (say, through 'inner perception') but fail to be secure by comparison to our access to other minds (say, both processes are equally susceptible to error).

The Peculiarity of Introspection

The peculiarity thesis maintains that self-ascriptions are arrived at differently, in an epistemically relevant way, from other-ascriptions. In order to determine how this may be so, we could start by characterising the way we arrive at other-ascriptions.

Third-person ascription is traditionally regarded as inferential and evidential (e.g., Ryle, 1949; Boghossian, 2003 [1989]). Moreover, the two most widely held views of the processes involved in third-person ascription, the Theory-Theory and the Simulation-Theory, involve processes which, although different, are both inferential and evidential (e.g., Davies and Stone, 1995a, 1995b; Carruthers and Smith, 1996). So, for example, the Theory-Theory maintains that we ascribe mental states to others by means of *Theory of Mind*-mediated inferences. The Theory of Mind is a rich body of data (either innate or acquired)¹⁴ that provides (at least) one of the premises required for the (tacit) inference. The inference involves two distinct types of premises: one type is the Theory of Mind-generalisation (for examples, see Churchland, 1981) and the other type is a premise about the particulars concerning the target. So, according to the Theory-Theory, the ascription of beliefs to others is reached through an inference from a generalisation and some particular evidence.

So, given third-person ascription is inferential and evidential, introspective self-ascription, negatively characterised, is to be a non-inferential, non-evidential process if the peculiarity thesis is to hold. Given this, we can appreciate that self-interpretative ascriptions of one's mental states do not count as special in the peculiarity sense. Since the ascription of mental states to others is inferential and evidential, the self-interpretative process does not differ from the third-person one. They are both inferential processes (although for all that has been said so far, distinct inferential processes might be involved) that employ behavioural and other kinds of evidence. This rules out an asymmetry between these two types of ascriptions of mental states.

So, although we can self-ascribe through self-interpretation, those ascriptions cannot be achieved differently. By contrast, introspective self-ascriptions seem intuitively to be arrived at differently.¹⁵ A *prima facie* reason for accepting this peculiarity thesis comes from the problem of other minds, which seems to stem from a fundamental asymmetry between our relation to our own and other minds. So assuming we would prefer to capture such peculiarity, next I introduce an introspective model that both is in line with our realist commitment and captures that peculiarity.

The Causal Model

The Causal Model of introspection maintains both that first-level states are *conceptually and ontologically independent* of the second-level states and that the former *cause* the latter. These are the two fundamental features of the model (cf. Shoemaker, 1994b). However, notice that the latter feature requires the former, given that in order for one state to cause the other, they need to be, as Hume says, "distinct existences" (ontologically distinct states).¹⁶ First of all, notice that these two

¹⁴ Some claim the theory develops via a general theory-forming capacity (Gopnik and Wellman, 1992). This process is meant to be analogous to the development of scientific theory. Others claim that the theory is innate (Carruthers, 1996). But these aren't the only options available. For example, one could, following Sterelny (2003), deny both of the above claims by holding that the generalisations are learned not through theorising, but through instruction, more particularly, by means of the epistemically engineered environment, which he calls "scaffolded learning."

¹⁵ Against the possibility of an illusion of peculiarity, as suggested by Gopnik (1993), see § 3.

¹⁶ So arguing against the ontological independence of the states (§ 8) will also count as an attack on the causal claim, since if the states were not distinct existences, they could not be related as cause and effect.

features of the model already allow us to accommodate realism about the mental (indeed, nothing is more real than what causes effects) and one of its main consequences: the fact that successful self-ascriptions involve a cognitive achievement (since they track their ontologically distinct self-ascribed states). Moreover, as one would expect, given that the connection between the first- and the second-level states is a contingent one, the model can easily make sense of self-ignorance and self-error (since the states are distinct existences, one could exist without the other).

So the Causal Model of introspection of beliefs claims that beliefs normally cause their self-ascriptions. But how do beliefs cause them? To this question different answers can be offered, given that there are different types of tracking. For example, the Inner Sense view holds that the introspective process is akin to perception (Locke, 1975, p. II.1.xix; Armstrong, 1993, p. 94). The “inner eye” detects by some sort of perception the belief to be self-ascribed. But this view seems wrong. There is no introspective sense-organ and introspection does not have a distinctive phenomenology, to mention two of the problems facing this view.¹⁷ Fortunately, the Causal Model can be divorced from the idea of inner perception. The tracking can exist without inner perception. The self-ascriptions can be caused by a non-perceptual mechanism, such as the one hypothesised by Nichols and Stich (2003): the *Monitoring Mechanism* (hereafter ‘MM’). In order to render the model more substantial, I will take MM to be the non-perceptual, introspective mechanism.¹⁸

Let me introduce MM by quoting from Nichols and Stich (2003, p. 160-161):

To have beliefs about one’s own beliefs, all that is required is that there be a [MM] that, when activated, takes the representation that p in the Belief Box as input and produces the representation *I believe that p* as output. This mechanism would be trivial to implement. To produce representations of one’s own beliefs, the [MM] merely has to copy representations from the Belief Box, embed the copies in a representation schema of the form: *I believe that ___*, and then place the new representations back in the Belief Box.

So MM takes beliefs as inputs and produces self-ascriptions as outputs, by embedding the copies of the beliefs in the schema *I believe that ___*. MM might seem simple, but that is no reason to think that it’s “trivial to implement,” since the beliefs that are being copied possess complex properties that give them the content and role they have in our mental economy. Although copying processes are common in organisms (for example, dividing cells copy their DNA each time they divide), whether this particular copying process will turn out to be trivial is not something we are currently in a position to assess. This, however, is no reason to reject the MM hypothesis.

So MM is taken as the introspective mechanism. However, the introspective process needs to be activated, as Nichols and Stich acknowledge. The input representation might just do that job, but then what is missing is a proposal for the input-feeding process. A fuller story will tell us just how such input is made available to MM.¹⁹ But this feeding process is meant to provide the input that activates

¹⁷ For other criticisms see, e.g., Shoemaker (1994a) and Cassam (2014). Of course, proponents of the view are aware of the difficulties generated by the analogy with perception and so try to play it down (e.g., Lycan, 1996, 2003), so much so that they are ‘Inner Sense’ theorists just in name.

¹⁸ However, notice the Causal Model does not require the adoption of the MM hypothesis. So the rejection of such hypothesis does not entail the rejection of the model.

¹⁹ Unfortunately I do not have a proposal for such process, but it will be a subpersonal process and probably the same process that allows us to feed beliefs to other systems. Consider the case where, after being asked to say something, you spontaneously say “I’m a human being,” and that it manifests a belief you hold. The idea is that whatever process allows you to feed such belief into the speech system might also be responsible for the feeding process of MM.

MM, which is responsible for the causal psychological transition from the belief to the self-ascription of the belief.

MM causally and contingently relates the belief and its self-ascription. Indeed, since the belief and its self-ascription are not conceptually and ontologically dependent and there is a connection between them, a naturalistic approach suggests that such connection is causal.²⁰ These features accommodate realism about the mental, make introspection a cognitive achievement and allow for the possibility of self-error and self-ignorance. All this makes a good *prima facie* case for the account and this support makes it compelling enough to be worth pursuing. But it only establishes that such an account is preferable to its competitors, all else being equal. But if this account had counterintuitive consequences, such as not capturing peculiarity, such support would probably not count for much. So, can the account accommodate the peculiarity of our introspective self-ascriptions?

I argue next that the account can capture the peculiarity thesis.²¹ Subsequently I defend the model against criticisms concerning the contingency to which it is committed.

The Causal Model and Peculiarity

The peculiarity thesis maintains that self-ascriptions are arrived at differently from other-ascriptions. The processes for arriving at first- and third-person mental ascriptions are to differ in kind.²² So, given that third-person ascriptions are based on inferences (either tacit or explicit) from evidence, the introspective process is, as mentioned, taken to be, negatively characterised, non-inferential and non-evidential. So the question we now need to answer is whether the Causal Model makes the process a non-inferential, non-evidential one.

But before considering that question, we need to make a clarification. The 'non-inferential, non-evidential process' constraint could be understood in two ways. One understanding of the constraint is that introspection cannot be a process that is *both* inferential *and* evidential. Another understanding is that introspection cannot be *either* an inferential *or* an evidential process. This is because it is at least possible to have an evidential process that is not inferential (e.g., Peacocke, 1999).²³ I examine here the Causal Model against the second reading of the constraint, given that this reading entails the first one. If the process does not end up being either inferential or evidential, the model captures (both readings of) the peculiarity claim.

So let's start by considering whether MM involves an evidential process. Is there, according to the Causal Model, any evidence which is gathered and on which the self-ascription is based? No. The self-ascription is based on the appropriate belief, but such relation is not evidential but brutally causal. Indeed, Peacocke (1999, p. 223ff) refers to models such as this as 'no-reasons' accounts and points out that the naturally accompanying epistemology of such accounts is reliabilism. This is because, given that

²⁰ This model is also friendly to epistemological naturalism, and particularly to reliabilism; see § 6.

²¹ This model is of course compatible with the security thesis. The introspective process could be more reliable than the one responsible for other-ascriptions.

²² However, it is important to keep in mind that the peculiarity thesis concerns only the introspective ascription of present mental states and the ascription of present mental states to others (i.e., the detection of mental states; see, e.g., Nichols and Stich, 2003, p. 151; Goldman, 2006, p. 223). So we are not concerned about the prediction of past and future mental states. It is anyway very likely that prediction in the first- and third-person cases is not different in kind.

²³ Peacocke argues that one's account needn't be inferential if one has reasons for the self-ascriptions (1999, p. 214-215). Sensations, such as pain, can be reasons for one's self-ascription, but they cannot be part of an inference (i.e., they cannot act as premises). This points to the possibility of a process being evidential but not inferential. Whether the self-ascription of beliefs is actually like that is a different matter.

the connection between the belief and the self-ascription is merely causal, what matters in the introspective process is that the causal connection is reliable. MM is fed a belief that is then copied and embedded in the appropriate schema. There is nothing involved in this process that deserves to be called 'evidential'. The process consists merely of contingent causal connections between beliefs and their self-ascriptions.

And does MM involve an inferential process? No. As characterised, MM involves a brutally causal process: a representation is copied and embedded in the appropriate schema. But the Causal Model seems compatible with the process being inferential. MM could resemble the Theory of Mind (if there is such a thing). But instead of comprising a network of generalisations, the body of data would consist of just one such generalisation: $p \rightarrow Bp$. So for every belief that p being fed to MM, one (tacitly) infers its self-ascription. Perhaps the right way to characterise such a process is as an inferentially thin process, as opposed to an inferentially thick one, but it might still be regarded as an inferential process, which would mean that this version cannot capture the peculiarity of self-ascription understood as a difference in kind.

But if 'inference' is not just any transition in thought (or processing of information, such as the causal one that the original MM version is meant to instantiate), how can this inference (deductive or inductive) be warranted? This is because, firstly, if $p \rightarrow Bp$ is thought of as a deductive inference, given that the truth of p does not entail the truth of Bp , the inference is invalid, hence devoid of warranting power. Secondly, if $p \rightarrow Bp$ is thought of as an inductive inference, given that such inference is false for most implications, this inference is also devoid of warranting power. So it looks as if this inference cannot yield any warrant.²⁴

Even if the move from p to Bp could be warranted, this view would still face difficulties. After all, if we had the generalisation $p \rightarrow Bp$, we would expect it to be just like other Theory of Mind-generalisations are meant to be. But, although we think that the third-person generalisations are true (all things considered), we don't think that $p \rightarrow Bp$ is. We do not think that if something is the case, then we believe it (in fact our implicit commitment to realism explains this). Indeed, we do not believe everything that is the case. Moreover, if the generalisations are acquired, through either theorising or, more plausibly, scaffolded learning (fn.12), then it is hard to imagine how we would actually acquire such a generalization, which seems so counterintuitive to us. So we can safely disregard such a characterisation of MM.

MM then is best understood as merely involving a causal process, as opposed to an inferential and evidential process. So the introspection of beliefs, according to the Causal Model, is a non-inferential, non-evidential process and is a different kind of process than the one employed in other-ascription of beliefs.²⁵ But some believe that the Causal Model cannot be faithful to the alleged intuition behind the peculiarity claim.

²⁴ The Displaced Perception model ("knowledge of internal (mental) facts via an awareness of external (physical) objects"; Dretske, 1995, p. 40—for more, see Byrne, 2005) tries to avoid these problems by requiring a connecting premise, but the move is ultimately unsuccessful (see Aydede, 2003).

²⁵ Now, if they are distinct processes, one would expect them to be subserved by numerically distinct mechanisms. So for the Causal Model to have empirical credibility, it is important that the evidence supports this (or, at least, does not stand against it). Unfortunately, the empirical evidence is inconclusive. It does not establish a double dissociation (for example, intact introspection and damaged other-ascription capacities in the case of autism, and damaged introspection and intact other-ascription capacities in the case of passivity forms of schizophrenia), but it is anyway compatible with it (see, e.g., Corcoran, 2000; Frith and Corcoran, 1996; Gerland, 1997; Grandin, 1984; Hobson, 2011; Hurlburt, 1990; Hurlburt *et al.*, 1994; Miedzianik, 1986; Parnas and Sass, 2011). So, importantly, the empirical evidence available does not rule a double dissociation out, thus not making the Causal Model an empirically unsound view (see also Carruthers, 2011). Anyhow, the mere fact that these two capacities are subserved by numerically distinct mechanisms does not by itself establish the peculiarity claim. The mechanisms need to be distinct *in kind*.

The Cartesian Intuition

It is sometimes believed that the peculiarity claim entails a principled asymmetry, which the Causal Model cannot accommodate. So Gertler claims that “if introspective states are merely causally related to introspected thoughts, the disparity between self-knowledge and other-knowledge is not truly principled” (2000, p. 125). Gertler takes seriously what she calls the Cartesian intuition about self-ascription: that is, the intuition that we have privileged access *in all possible worlds*.

It is easy to appreciate why the Causal Model does not seem to do justice to this intuition: it seems possible, according to this view, to have the same kind of access to someone else’s mind. In some possible world, a mechanism that, like MM, ascribes (non-inferentially and non-evidentially) mental states might be connected via “special rewiring” (Lycan, 1996, p. 49) to another person.²⁶ So, given the Causal Model, the peculiarity of self-ascriptions won’t hold in every possible world.

But how much of a problem is this? Well, it is not clear that most of us have any such intuition. Some philosophers (such as Tyler Burge, Stephen Wright and Sidney Shoemaker) might share this intuition (though notice that Gertler (2000, p. 126) thinks that the Cartesian intuition is limited to occurrent phenomena, not dispositional ones), but other philosophers and laypeople do not. Indeed, David Armstrong (Armstrong and Malcolm, 1984) thinks that rejecting this intuition is an asset since he takes “the claim that telepathic knowledge [is possible] to be the claim that we [can] have some direct awareness of the mental states of others” (Armstrong, 1993, p. 124).²⁷ Given that we can conceive this possibility, the asymmetry needn’t hold in all possible worlds. So there is no reason to require a principled disparity between self- and other-ascription, which means that the Causal Model can capture the peculiarity thesis by being a non-inferential, non-evidential causal process.

This concludes my attempt to show that the Causal Model accounts for the peculiarity of self-ascription. Along the way, we pre-empted objections against the model but there is still one main objection that we have not faced. In the next section, I introduce the objection and argue that it is unsuccessful.

The Contingency of Self-Ascriptions

Some philosophers believe that introspective self-ascriptions cannot be contingently related to their subject matter. Such contingency is attacked more fervently by Sidney Shoemaker. Shoemaker argues that there is a necessary link between introspection and rationality, intelligence and conceptual capacities. So let’s consider his case.

Shoemaker (1994b, 1994c [1988]) puts forward the *argument from self-blindness*. This argument poses a direct threat to the Causal Model because it attacks views where the relation between beliefs and their self-ascriptions is contingent. The argument goes as follows: (1) if the connection between beliefs and their

²⁶ So this mechanism is fed the representations from the other’s belief box, embeds them in a schema of the form *S/he believes that* ___ and such embedded representations are placed in one’s belief box.

²⁷ Interestingly, in some disorders of the self, “the subject feels that her own thought is shared by others, broadcast without her permission or withdrawn by others” (Bortolotti, 2010, p. 228). Consider also cases of thought ‘insertion’, where a subject is under the impression of having direct access to someone else’s thoughts (“the thoughts of Eamonn Andrews come into my mind. There are no other thoughts there, only his [...]. He treats my mind like a screen and flashes thoughts onto it like you flash a picture”—Mellor, 1970, p. 17; “Thoughts come into my head like ‘Kill God’. It is just like my mind working, but it is not. They come from this chap, Chris. They’re his thoughts”—Frith, 1992, p. 66).

self-ascriptions is contingent, then self-blindness is a conceptual possibility, but (2) self-blindness is not a conceptual possibility, so (3) the connection between beliefs and their self-ascriptions is not contingent. Before we can judge (1) and (2), we need to introduce the pivotal notion of 'self-blindness.' A subject is self-blind when she lacks *introspective* awareness of her beliefs even while possessing normal intelligence, rationality and conceptual capacities. So self-blindness does not entail any cognitive deficiency, just like the "person who lacks sight can in principle be equal in intelligence, rationality, and conceptual capacity to any sighted person" (1994b, p. 236). And although the subject cannot have self-knowledge through introspection, it can achieve such knowledge through self-interpretation: that is, through reasoning, in an inferential and evidential way (Shoemaker, 1994c, p. 30-31).

Premise (1) seems uncontroversial. Given that beliefs and their self-ascriptions exist independently, it is possible for a belief never to be introspected. Indeed, it is possible that none of our beliefs be introspected. The contingency of self-ascriptions allows for the possibility of self-blindness. So given that the argument is valid, if we are to defend the contingency of self-ascriptions, our only alternative is to find premise (2) wanting. Shoemaker is aware of the fact that (2) is the key premise of his argument and so puts forward two main arguments to support it: the *argument from Moore's Paradox* and the *argument from the utility of introspection*.²⁸ Let's take them in turn.

The argument from Moore's Paradox (1994c, p. 34-37) is intended as a *reductio* of the possibility of self-blindness. Shoemaker's strategy is to first consider what a self-blind person would be like and then argue that, given normal intelligence, rationality and conceptual capacities, such a person could not exist. Let me expand.

Shoemaker claims that it is possible for the self-blind person to have evidence for *p*, while the "third-person" evidence available to him supports the proposition that he does not believe that *p* (1994c, p. 35). So, if this were to happen, the self-blind person would be susceptible to uttering a Moorean sentence of the form '*p*, but I don't believe that *p*'. But Shoemaker claims that, given that the self-blind person has normal intelligence, rationality and conceptual capacities (especially concerning the concept of belief), the self-blind person would be as capable as we are of recognising the paradoxical character of Moorean sentences and hence avoid them (1994c, p. 35). So, contrary to what we first thought, the self-blind person is not susceptible to making Moorean assertions. But now the self-blind person "is beginning to look just like a normal person. It would appear that there would be nothing in his behaviour, verbal or otherwise, that would give away the fact that he lacks [introspection]" (1994c, p. 36). And Shoemaker thinks that if one behaves *as if* one has introspective awareness of beliefs, then one does have such awareness (1994c, p. 34, 36). So, according to Shoemaker, the self-blind person has introspective awareness of his beliefs, contrary to what *ex-hypothesi* is meant to be the case, hence the *reductio*. This is meant to show then the impossibility of self-blindness: that is, the truth of (2). But is this argument successful? There are at least two reasons to believe that it is not.

First, although the self-blind person begins to look just like a normal person, it is not clear their behaviour is equivalent. Given that the self-blind person's self-ascription is based on third-person evidence from which he reasons, if he was asked

²⁸ Given I am interested in the introspection of *beliefs*, I am ignoring other arguments Shoemaker puts forward for the impossibility of the lack of introspective awareness of different occurrent phenomena, such as pain and perceptual experiences (1994b, p. 226-231). I am ignoring them not because the arguments are not directly intended to apply to beliefs, but because they cannot be extended to apply to them given that the arguments rely on the different nature of the phenomena to which they are intended to apply.

how he arrived at such self-ascription or the grounds for such self-ascription, one would expect a different answer from the one a normal (non-self-blind) person would give.²⁹ That is a potential behavioural mismatch, which mirrors the different processes involved in the self-ascription of the self-blind and the normal persons.

Second, Shoemaker relies on the following principle: one's behaviour provides the best possible evidence that one is introspectively aware of one's beliefs (1994c, p. 34-36, 1994b, p. 236). The thing to note about it is that its best chance to be true is if applied broadly to awareness of one's beliefs, rather than *introspective* awareness (which explains why Shoemaker sometimes seems to equivocate between the two notions when talking about this principle; see, e.g., 1994c, p. 34, 1994b, p. 236). Let me explain.

Given normal intelligence, rationality and conceptual capacities, it is easy to see that a person who only possesses third-person evidence can work out what they believe and even avoid Moorean assertions. So a behavioural replica of a subject, with normal intelligence, rationality and conceptual capacities, seems to also have awareness of his mind. But this falls short of showing that such replica would also have *introspective* access to his mind. Unless one rules out the possibility that such self-awareness is achieved through self-interpretation, one cannot claim to have shown that such awareness is introspective. As Amy Kind (2003, p. 45) rightly says:

At least in principle, [the fact that George, the allegedly self-blind person, is aware of his own beliefs to the same extent as a normal person would be] is consistent with the claim that George *is* self-blind, but has acquired the normal extent of awareness of his own beliefs [...] via third-person access to these states. Self-blindness is defined in terms of [introspection], not [self-awareness].

Perhaps one only needs to show that such self-awareness is not entirely achieved through self-interpretation, so that one could claim that we have at least some introspective access to our minds. One might here attempt to apply Boghossian's insights about the limits of inferential and evidential accounts of self-awareness (§ 3). But, as noted, such accounts face problems when trying to account for one's awareness of one's mind when there seems to be no available evidence to infer its contents. But these are *not* the cases that Shoemaker has in mind when describing how George has, after all, awareness of his mind and so he starts to look just like a normal (non-self-blind) person. The cases Shoemaker focuses on are cases where there *is* evidence for George to infer that he believes so-and-so, and this is how Shoemaker can claim that George and the normal person are beginning to look like each other.

So depending on whether George has the evidence or not, *either* Shoemaker cannot show that George does not lack introspective awareness (since his awareness can be accounted for by means of self-interpretation) *or* he fails to show that George is like the normal person (since the evidence is what allows the similarity claim). Either way, Shoemaker's *reductio* does not go through. The problem, in a nutshell, is that Shoemaker says nothing to rule out the possibility of the self-blind person's self-awareness being achieved through self-interpretation (exclusively), and he cannot help himself to Boghossian's insights about inferential and evidential accounts of self-ascription since his claim (required for the *reductio*) that the

²⁹ In fact, only in the case of the person with introspective access the question, say, "How do you know that you believe that Elvis was really blonde?" seems an unreasonable one. You just do (see, e.g., Neta, 2011, p. 18-19): there is no reason that supports your introspective self-ascription. This is what makes such self-ascriptions *groundlessly* warranted.

self-blind and the normal persons are behaviourally alike is based on the fact that there is evidence for the self-blind person self-ascribing beliefs. So Shoemaker is not entitled to the claim that self-blindness is impossible on the basis of the argument from Moore's Paradox.

So, what about the argument from the utility of introspection? This argument (1994c, p. 27-30; 1994b, p. 237-240) is really based on a pair of considerations about the utility of self-knowledge, rather than introspection, and so it fails to show the impossibility of self-blindness for basically the same reasons the previous one does. Let me explain.

Shoemaker points out that self-knowledge plays two useful roles in our lives: one in achieving our aims through cooperation and another in deliberation (1994c, p. 27-30). With respect to the first one, Shoemaker claims that a 'rational agent' (i.e., a person with normal intelligence, rationality and conceptual capacity) will realise the benefits of cooperation from others in order to achieve one's aims and that getting this cooperation involves making clear what one believes, including saying 'I believe that...' (1994b, p. 238).³⁰ So, again, the rational agent, who *ex-hypothesi* lacks introspective access to his beliefs, will act as if he has such access. And, again, given the above principle (if someone acts as if she has introspective access, she has such access), Shoemaker can claim that, given intelligence, rationality and conceptual capacity, we do have introspective access to our beliefs. The problem, however, is that (regardless of the dubious nature of the principle), again, Shoemaker does not rule out the possibility of such self-knowledge being achieved through self-interpretation. Unless one can show that such self-knowledge cannot be achieved through self-interpretation, we do not have an argument against the possibility of self-blindness.

Shoemaker might at this point reply that, from an evolutionary perspective,

[It] would certainly be bizarre to suppose that, having endowed creatures with everything necessary to give them a certain very useful behavioural repertoire – namely that of creatures with normal intelligence, rationality, and conceptual capacity, plus the ability to acquire first order beliefs about the environment from sense-perception – Mother Nature went to the trouble of installing in them an *additional* mechanism, a faculty of Inner Sense [or the MM], whose impact on behaviour is completely redundant, since its behavioural effects are ones that would occur anyhow as the result of the initial endowment (1994b, p. 239-240).

However, as noted, it is not at all clear that the additional mechanism would not have an impact on behaviour. For example, the additional mechanism could make such self-ascriptions much faster and therefore bring evolutionary advantages to the creatures endowed with it. Furthermore, such mechanism would dispense with reasoning, thus allowing for parallel processing and leaving the reasoning capacity available for other important issues for which no mechanism can be selected (say, due to the unstable nature of the subject-matter). So the additional mechanism's impact wouldn't be redundant and, from an evolutionary perspective, there would be reasons for such mechanism to be selected.³¹

With respect to the second useful role that self-knowledge plays in our lives, Shoemaker claims that it is essential to being rational that one engages in delibera-

³⁰ Note that, for social non-human animals that engage in cooperation, this does not seem necessary; however, they are not thought to be 'rational agents.'

³¹ Just as there seems to be reasons for thinking that two systems of reasoning have evolved, System 1 and System 2: S1 mostly operates automatically and quickly, with little or no effort, and no sense of voluntary control, and often at a subpersonal level, whereas S2 is slow but deliberate, orderly and effortful (see, e.g., Evans, 2007; Evans and Over, 1996; Kahneman, 2011; Stanovich, 2004; see also Kornblith, 2012).

tion (1994c, p. 28-29; 1994b, p. 240). Now, regardless of whether we understand this claim in terms of reflective maintenance (i.e., the reviewing of beliefs) or in terms of what Tyler Burge (1996) calls critical reasoning (i.e., reasoning about reasons as reasons), the argument does not go through. Even if we agree that both sorts of reasoning require self-knowledge (but see Owens, 2011, p. 269; Cassam, 2014, p. 41-42, regarding the need of self-knowledge for reasoning), that falls short of showing that they require introspection. Again, unless the possibility of such knowledge being achieved through self-interpretation is ruled out, Shoemaker does not have a case for a necessary link between rationality, intelligence and conceptual capacity, on the one hand, and introspection, on the other. For this, one would have to show that reasons cannot (always) be known through self-interpretation, and Shoemaker does not show that.

But, anyway, even if one assumes that either reasoning requires introspection, this needn't mean that lack of introspection is impossible. It can mean that self-blindness has some higher-cognitive consequences, contrary to what Shoemaker claims. The parallel between (non-native) blindness and self-blindness might be misplaced (1994b, p. 236).³² In fact, one would expect introspection to have such higher-cognitive consequences, given that one is now taking rationality to presuppose introspection.³³

Moreover, although our introspective capacity might be required for higher types of reasoning, this does not entail that if one engages in such reasoning then our capacity for introspection comes for free with such abilities, intelligence and so on (as Shoemaker believes, 1994b, p. 243). It only entails that we do have such a capacity, and the Causal Model here suggested sustains that such capacity is subserved by MM.

Also, and keeping in mind that not every kind of deliberation requires introspection (Owens, 2011), notice that the above notion of rationality trivialises the claim that self-blindness is impossible. That is, if deliberation requires introspection and the requirement for rationality includes being a deliberator, and given that the self-blind person is *ex-hypothesi* rational, self-blindness is trivially (i.e., by definition) impossible. More importantly, there does not seem to be any reasons for accepting that the self-blind person is rational in this sense.

So I conclude that both of Shoemaker's arguments that attempt to establish the impossibility of self-blindness, premise (2), are unsuccessful. Consequently, Shoemaker's argument from self-blindness isn't cogent, given that there is no reason to hold the key premise (2) true, and so he fails to argue against the contingency of self-ascriptions. Shoemaker offers us no reason to reject the Causal Model on the basis that it posits a contingent connection between beliefs and their self-ascriptions. In fact, we have already seen (§ 5) that there is a strong *prima facie* case for such contingency given a realist commitment.

Concluding remarks

Some might anyway complain at this point that, despite these and other advantages, this project was not worth pursuing from the outset, given that it does not aim to give an account of self-ascription of mental states in general (e.g., Cassam, 2014; Goldman, 2006). That is, a single account that explains all self-ascription.

³² A perceptual account of introspection might foster such a mistake.

³³ Just as one would expect higher-cognitive consequences if one had a damaged language (modular) faculty *a la* Chomsky (see, e.g., Smith, 1999).

Comprehensive, and so more parsimonious, accounts are normally desirable and often desired. But this degree of simplicity and elegance cannot be achieved when considering the general phenomenon of self-ascription. Although these theoretical virtues should and do influence our intellectual practices, in the case at hand they can only shape the investigation at the expense of inadequacy.

Acknowledging the distinction between occurrent and dispositional mental phenomena represents only the beginning of a proper investigation. Already with this distinction in place, we can see that general accounts of self-ascription that exploit properties that are not shared by all mental phenomena will not succeed (such as Goldman's Phenomenological account; see § 2, cf. Goldman 2006). To make matters worse, claims about privileged access vary depending on the different natures of these phenomena. So, for example, self-ascriptions of sensations (such as pain) are often thought of as being more resilient to error (Gertler, 2008; see also Wright, 1998). This, however, does not mean that propositional attitudes could not be as secure as sensations are. If Burge (1996) is correct, *cogito*-like thoughts are self-verifying and so infallible. But it is likely that propositional attitudes that are dispositional, such as beliefs, do not enjoy this greater security.

So it seems that not even a general account of self-ascription for the propositional attitudes is to be expected (as some appreciate; e.g., Moran, 2001). The diverse natures of the different propositional attitudes seem to prevent such an account. Indeed, even within a more constrained folk-psychological category we might find fundamental differences. For instance, if we distinguish between higher-cognitive emotions, such as romantic love and jealousy, and basic emotions, such as fear and anger (see, e.g., Griffiths, 1997), we can appreciate that the account of self-ascription of one set of phenomena and of the other might easily differ. Self-ascription of higher-cognitive emotions is likely to be only the product of self-interpretation (say, due to their more recent history in the species and the frequent self-deception that accompanies them), whereas self-ascription of basic emotions might be also the product of introspection (say, due to the clear evolutionary advantage that this would provide and universality). Indeed, a mechanism very much like MM could perhaps subserve such introspective capacity. Perhaps even dispositional states other than beliefs could be subserved by (distinct) MM mechanisms (see Nichols and Stich, 2003). But it definitely remains to be seen whether the Causal Model can plausibly be extended to other mental states.

However, the point stands: beliefs, sensations and passing thoughts (to mention some) are all very different mental phenomena and there is no reason to think that a general account of self-ascription can be achieved. It is extremely likely that a plurality of processes are required to explain our capacity for self-ascribing (as some appreciate; e.g., Schwitzgebel, 2012), provided one does not overlook the differences between distinct types of mental states or does not focus on a limited range of phenomena. And one such process is the non-inferential, non-evidential, causal process subserved by MM for the introspection of beliefs.

So to conclude, I have developed and defended an account of the introspection of beliefs: the Causal Model. We have seen that the main thesis of the model is that the tracking of a belief is causal and contingent. I have, to this end, put forward a specific mechanism to flesh out the proposal: MM. This mechanism subserves the introspective process. This process is different in kind from the process of ascribing beliefs to others.

Moreover, I have tackled along the way a series of criticisms. We have, for example, seen that the usual problems that Inner Sense accounts of introspection face do not affect this account. But perhaps more importantly, given that the Causal Model is committed to a contingent relation between beliefs and their

self-ascriptions, we have seen that attacks on such a connection are not successful. Indeed, we have seen that there is a *prima facie* case for this contingency, which in turn points to more assets of this account, such as rendering self-ascription a genuine cognitive achievement and straightforwardly making sense of a variety of phenomena, including self-error and self-ignorance. In fact, the Causal Model captures all of the constraints set by our starting point: realism about the mental. All this shows that the Causal Model is, from a realist (and naturalist) perspective, an excellent candidate for explaining our peculiar introspective access to beliefs.

The Causal Model, supplemented by the MM hypothesis, is then the explanation I propose for how we access in a special way our own beliefs. The Causal Model does have its more than fair share of critics and indeed the model has, in recent years, fallen out of favour in the philosophical world. However, I hope this paper has shown that it should not have.

References

- ALSTON, W. 1971. Varieties of Privileged Access. *American Philosophical Quarterly*, 8:223-41.
- ARMSTRONG, D.M. 1993. *A Materialist Theory of the Mind*. London, Routledge, 375 p.
- ARMSTRONG, D.M.; MALCOM, N. 1984. *Consciousness and Causality: A Debate on the Nature of Mind*. Oxford, Blackwell, 222 p.
- AYDEDE, M. 2003. Is Introspection Inferential? In: B. GERTLER (ed.), *Privileged Access*. Hants, Ashgate, p. 55-64.
- BAR-ON, D. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford, Oxford University Press, 449 p. <http://dx.doi.org/10.1093/0199276285.001.0001>
- BOGHOSSIAN, P.A. 2003 [1989]. Content and Self-Knowledge. In: B. GERTLER (ed.), *Privileged Access*. Hants, Ashgate, p. 65-82.
- BORTOLOTTI, L. 2010. *Delusions and Other Irrational Beliefs*. Oxford, Oxford University Press, 299 p.
- BURGE, T. 1996. Our Entitlement to Self-Knowledge. *Proceedings of the Aristotelian Society*, **XCVI**:91-116. <http://dx.doi.org/10.1093/aristotelian/96.1.91>
- BYRNE, A. 2005. Introspection. *Philosophical Topics*, **33**(1):79-104. <http://dx.doi.org/10.5840/philtopics20053312>
- CARRUTHERS, P. 1996. Simulation and Self-Knowledge: A Defence of Theory-Theory. In: P. CARRUTHERS; P. SMITH (eds.), *Theories of Theories of Mind*. Cambridge, Cambridge University Press, p. 22-38. <http://dx.doi.org/10.1017/CBO9780511597985.004>
- CARRUTHERS, P. 2011. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford, Oxford University Press, 437 p. <http://dx.doi.org/10.1093/acprof:oso/9780199596195.001.0001>
- CARRUTHERS, P.; SMITH, P. 1996. *Theories of Theories of Mind*. Cambridge, Cambridge University Press, 390 p. <http://dx.doi.org/10.1017/CBO9780511597985>
- CASSAM, Q. 2014. *Self-Knowledge for Humans*. Oxford, Oxford University Press, 238 p.
- CHURCHLAND, P.M. 1981. Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy*, **78**:67-90. <http://dx.doi.org/10.2307/2025900>
- CHURCHLAND, P.M. 1992. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MIT Press, 321 p.
- CORCORAN, R. 2000. Theory of Mind in Other Clinical Conditions: Is a Selective 'Theory of Mind' Deficit Exclusive to Autism? In: S. BARON-COHEN; H. TAGER-FLUSSBERG; D.J. COHEN (eds.), *Understanding Other Minds*. 2nd ed., Oxford, Oxford University Press, p. 391-421.
- CRANE, T. 2001. *Elements of Mind: An Introduction to the Philosophy of Mind*. Oxford, Oxford University Press, 181 p.
- DAVIES, M.; STONE, M. 1995a. *Folk Psychology: The Theory of Mind Debate*. Oxford, Blackwell, 301 p.
- DAVIES, M.; STONE, M. 1995b. *Mental Simulation: Evaluations and Applications*. Oxford, Blackwell, 286 p.
- DENNETT, D. 1987. *The Intentional Stance*. Cambridge, MIT Press, 388 p.

- DRETSKE, F. 1995. *Naturalizing the Mind*. Cambridge, MIT Press, 208 p.
- DESCARTES, R. 1984. *The Philosophical Writings of Descartes. Vol. I & II*. Cambridge, Cambridge University Press, 845 p.
- DUMMETT, M. 1978. *Truth and Other Enigmas*. London, Duckworth, 470 p.
- EVANS, J. 2007. *Hypothetical Thinking: Dual Processes in Reasoning and Judgment*. Hove, Psychology Press, 206 p.
- EVANS, J.; OVER, D. 1996. *Rationality and Reasoning*. Hove, Psychology Press, 179 p.
- FERNANDEZ, J. 2003. Privileged Access Naturalized. *The Philosophical Quarterly*, 53:352-372. <http://dx.doi.org/10.1111/1467-9213.00317>
- FERNANDEZ, J. 2013. *Transparent Minds: A Study of Self-Knowledge*. Oxford, Oxford University Press, 245 p.
<http://dx.doi.org/10.1093/acprof:oso/9780199664023.001.0001>
- FRICKER, E. 1998. Self-Knowledge: Special Access versus Artefact of Grammar – A Dichotomy Rejected. In: C. WRIGHT; B.C. SMITH; C. MACDONALD (eds.), *Knowing Our Own Minds*. Oxford, Oxford University Press, p. 155-206.
- FRITH, C. 1992. *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, Erlbaum, 169 p.
- FRITH, C.; CORCORAN, R. 1996. Exploring 'Theory of Mind' in People with Schizophrenia. *Psychological Medicine*, 26:521-30.
<http://dx.doi.org/10.1017/S0033291700035601>
- GALLOIS, A. 1996. *The World without, the Mind within: An Essay on First-Person Authority*. Cambridge, Cambridge University Press, 213 p.
<http://dx.doi.org/10.1017/CBO9780511627880>
- GERLAND, G. 1997. *A Real Person: Life on the Outside*. London, Souvenir, 258 p.
- GERTLER, B. 2000. The Mechanics of Self-Knowledge. *Philosophical Topics*, 28:125-46. <http://dx.doi.org/10.5840/philtopics200028215>
- GERTLER, B. 2003. *Privileged Access: Philosophical Accounts of Self-Knowledge*. Hants, Ashgate, 266 p.
- GERTLER, B. 2008. Self-Knowledge. In: ZALTA (ed.), *The Stanford Encyclopedia of Philosophy*. Available at: <http://plato.stanford.edu/entries/self-knowledge/> Accessed on: 02/20/2015.
- GERTLER, B. 2010. *Self-Knowledge*. London, Routledge, 314 p.
- GOLDMAN, A.I. 1993. The Psychology of Folk Psychology. *Behavioural and Brain Sciences*, 16:15-28. <http://dx.doi.org/10.1017/S0140525X00028648>
- GOLDMAN, A.I. 2006. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford, Oxford University Press, 364 p.
<http://dx.doi.org/10.1093/0195138929.001.0001>
- GOPNIK, A. 1993. How We Know Our Minds: The Illusion of First-Person Knowledge of Intentionality. *Behavioural and Brain Sciences*, 16:1-14.
<http://dx.doi.org/10.1017/S0140525X00028636>
- GOPNIK, A.; WELLMAN, H. 1992. Why the Child's Theory of Mind is Really a Theory. *Mind and Language*, 7:145-72.
<http://dx.doi.org/10.1111/j.1468-0017.1992.tb00202.x>
- GRANDIN, T. 1984. My Experiences as an Autistic Child and Review of Selected Literature. *Journal of Orthomolecular Psychiatry*, 13:144-175.
- GRECO, J. 2010. *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge, Cambridge University Press, 205 p.
<http://dx.doi.org/10.1017/CBO9780511844645>
- GRIFFITHS, P. 1997. *What Emotions Really Are*. Chicago, Chicago University Press, 302 p. <http://dx.doi.org/10.7208/chicago/9780226308760.001.0001>
- HOBSON, P. 2011. Autism and the Self. In: S. GALLAGHER (ed.), *The Oxford Handbook of the Self*. Oxford, Oxford University Press, p. 571-591.
<http://dx.doi.org/10.1093/oxfordhb/9780199548019.003.0025>
- HURLBURT, R.T. 1990. *Sampling Normal and Schizophrenic Inner Experience*. New York, Plenum Press, 287 p. <http://dx.doi.org/10.1007/978-1-4757-0289-7>
- HURLBURT, R.T.; HAPPE, F.; FRITH, U. 1994. Sampling the Form of Inner Experience in Three Adults with Asperger Syndrome. *Psychological Medicine*, 24:385-395.
<http://dx.doi.org/10.1017/S0033291700027367>
- KAHNEMAN, D. 2011. *Thinking, Fast and Slow*. London, Allen Lane, 499 p.
- KIND, A. 2003. Shoemaker, Self-Blindness and Moore's Paradox. *The Philosophical Quarterly*, 53:39-48. <http://dx.doi.org/10.1111/1467-9213.00294>
- KLAUSEN, S. 2008. The Phenomenology of Propositional Attitudes. *Phenomenology and the Cognitive Sciences*, 7:445-462.
<http://dx.doi.org/10.1007/s11097-007-9081-z>

- KORNBLITH, H. 2012. *On Reflection*. Oxford, Oxford University Press, 188 p.
<http://dx.doi.org/10.1093/acprof:oso/9780199563005.001.0001>
- LOCKE, J. 1975. *An Essay Concerning Human Understanding*. Oxford, Oxford University Press, 748 p.
- LYCAN, W. 1996. *Consciousness and Experience*. Cambridge, MIT Press, 211 p.
- LYCAN, W. 2003. Dretske's Ways of Introspecting. In: B. GERTLER (ed.), *Privileged Access*. Hants, Ashgate, p. 15-29.
- MELLOR, C. 1970. First Rank Symptoms of Schizophrenia. I. The Frequency of Schizophrenics on Admission to Hospital. II. Differences between Individual First Rank Symptoms. *British Journal of Psychiatry*, 117:15-23.
- MIEDZIANIK, D. 1986. *My Autobiography*. Nottingham, Nottingham University Press, 115 p.
- MORAN, R. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. New Jersey, Princeton University Press, 202 p.
- NETA, R. 2011. The Nature and Reach of Privileged Access. In: A. HATZIMOYSIS (ed.), *Self-Knowledge*. Oxford, Oxford University Press, p. 9-32.
<http://dx.doi.org/10.1093/acprof:oso/9780199590728.003.0002>
- NICHOLS, S.; STICH, S. 2003. *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford, Oxford University Press, 237 p. <http://dx.doi.org/10.1093/0198236107.001.0001>
- OWENS, D. 2011. Deliberation and the First Person. In: A. HATZIMOYSIS (ed.), *Self-Knowledge*. Oxford, Oxford University Press, p. 261-278.
<http://dx.doi.org/10.1093/acprof:oso/9780199590728.003.0014>
- PARNAS, J.; SASS, L. 2011. The Structure of Self-Consciousness in Schizophrenia. In: S. GALLAGHER (ed.), *The Oxford Handbook of the Self*. Oxford, Oxford University Press, p. 521-546.
<http://dx.doi.org/10.1093/oxfordhb/9780199548019.003.0023>
- PEACOCKE, C. 1999. *Being Known*. Oxford, Oxford University Press, 368 p.
<http://dx.doi.org/10.1093/0198238606.001.0001>
- RYLE, G. 1949. *The Concept of Mind*. London, Penguin Books, 282 p.
- SCHWITZGEBEL, E. 2012. Introspection, What. In: D. SMITHIES; D. STOLJAR (eds.), *Introspection and Consciousness*. Oxford, Oxford University Press, p. 29-48.
<http://dx.doi.org/10.1093/acprof:oso/9780199744794.003.0001>
- SHOEMAKER, S. 1994a. Self-Knowledge and "Inner Sense." Lecture I: The Object Perception Model. In: S. SHOEMAKER, *The First-Person Perspective and Other Essays*. Cambridge, Cambridge University Press, p. 201-223.
<http://dx.doi.org/10.2307/2108488>
- SHOEMAKER, S. 1994b. Self-Knowledge and "Inner Sense." Lecture II: The Broad Perceptual Model. In: S. SHOEMAKER, *The First-Person Perspective and Other Essays*. Cambridge, Cambridge University Press, p. 224-245.
<http://dx.doi.org/10.2307/2108489>
- SHOEMAKER, S. 1994c. On Knowing One's Own Mind. In: S. SHOEMAKER, *The First-Person Perspective and Other Essays*. Cambridge, Cambridge University Press, p. 25-49.
- SHOEMAKER, S. 1994d. *The First-Person Perspective and Other Essays*. Cambridge, Cambridge University Press, 278 p.
- SMITH, B.C. 1998. On Knowing One's Own Language. In: C. WRIGHT; B.C. SMITH; C. MACDONALD (eds.), *Knowing Our Own Minds*. Oxford, Oxford University Press, p. 391-428.
- SMITH, N. 1999. *Chomsky: Ideas and Ideals*. Cambridge, Cambridge University Press, 268 p. <http://dx.doi.org/10.1017/CBO9781139163897>
- SMITHIES, D.; STOLJAR, D. 2012. *Introspection and Consciousness*. Oxford, Oxford University Press, 448 p.
<http://dx.doi.org/10.1093/acprof:oso/9780199744794.001.0001>
- SOSA, E. 2007. *Apt Belief and Reflective Knowledge, Volume 1: A Virtue Epistemology*. Oxford, Oxford University Press, 149 p.
<http://dx.doi.org/10.1093/acprof:oso/9780199297023.001.0001>
- STANOVICH, K. 2004. *The Robot's Rebellion: Finding Meaning in the Age of Darwin*. Chicago, University of Chicago Press, 358 p.
<http://dx.doi.org/10.7208/chicago/9780226771199.001.0001>
- STERELNY, K. 2003. *Thought in a Hostile World: The Evolution of Human Cognition*. Oxford, Blackwell, 262 p.
- STROUD, B. 2000. *The Quest for Reality: Subjectivism and the Metaphysics of Colour*. New York, Oxford University Press, 228 p.

- TANNEY, J. 2002. Self-Knowledge, Normativity, and Construction. *In*: A. O'HEAR (ed.), *Logic, Thought and Language*. Cambridge, Cambridge University Press, p. 37-55. (Royal Institute of Philosophy Supplement, 51).
- WITTGENSTEIN, L. 1953. *Philosophical Investigations*. Oxford, Blackwell, 232 p.
- WOLLHEIM, R. 1999. *On The Emotions*. New Haven, Yale University Press, 269 p.
- WRIGHT, C. 1998. Self-Knowledge: The Wittgensteinian Legacy. *In*: C. WRIGHT; B.C. SMITH; C. MACDONALD (eds.), *Knowing Our Own Minds*. Oxford, Oxford University Press, p. 13-46. <http://dx.doi.org/10.1017/cbo9780511563744.007>

Submitted on March 24, 2015

Accepted on October 5, 2015