



# The Sense of Agency and the Epistemology of Thinking

Casey Doyle<sup>1</sup>

Received: 10 September 2019 / Accepted: 24 August 2020  
© Springer Nature B.V. 2020

## Abstract

This paper motivates a constraint on how to explain the “sense of agency” for conscious thinking. It argues that a prominent model fails to satisfy the constraint before sketching an alternative that does. On the alternative, punctate acts of conscious thinking, such as episodes of inner speech, are recognizable as our deeds because they are recognizable as parts of complex cognitive activities, which we know non-observationally in virtue of holding intentions to perform them.

## 1 Introduction

In the parlance, the “sense of ownership” refers to the phenomenon of taking one’s bodily parts and movements or mental states as one’s own. The “sense of agency” refers to the phenomenon of taking one’s bodily movements or mental acts as one’s own doings.<sup>1</sup> Each sense divides into two types, according to its object: senses of *bodily* ownership and agency and senses of *psychological* ownership and agency. This paper concerns only one of these: the sense of psychological agency, specifically the sense of agency over one’s own acts of conscious thinking. The sense of agency under discussion here is the phenomenon of taking some stretch of conscious thinking as one’s own doing.

Thinking is a broad category, encompassing such activities as imagining, reasoning, comparing, planning, searching one’s memory, and the like. “Conscious thinking” refers to any such activity that is realized in conscious experience. Perhaps not all thinking is conscious and not all parts of a single episode of thinking need be conscious. For example, you might plan a vacation off and on over the course of a week, but only some parts of that are manifest in conscious experience. Because the phenomena are diverse, it helps to focus on a paradigmatic case: inner speech.

<sup>1</sup> The literature on the senses of ownership and agency is extensive. For ownership see Martin (1995), the essays collected in Roessler and Eilan (2003), Zahavi (2005), and Musholt (2015). For the sense of agency, in addition to the essays discussed below, see Bayne (2008).

✉ Casey Doyle  
caseyb.doyle@gmail.com

<sup>1</sup> Binghamton University, Binghamton, New York, USA

This is a process of auditory imagining that unfolds over time, that occupies one's attention, and that is explicitly *there* before the mind's eye (or ear). As Hulburt et al. put it, "the speakings are generally apprehended to be in the person's own naturally inflected voice, in the same rhythm, pacing, expressivity, tone, hesitations, and style as external speaking" (2013, 1482).

Not all conscious thinking is realized in inner speech, of course. One reason to take inner speech as a paradigm is that it occurs frequently and effortlessly, and when it occurs it is, for the most part, obvious to the subject. It is what Hulburt and Heavey call a "pristine inner experience" (2018, 168). Still, there is considerable controversy about how to understand inner speech. I assume that inner speech involves an auditory-phonological component, but remain neutral on whether it has this essentially (Langland-Hassan 2018; Wilkinson and Ferynhough 2018) or as an additional component to a thought tokening (Gauker 2018). I assume also that inner speech is a medium in which we can engage in spontaneous thinking.<sup>2</sup> Furthermore, inner speech is itself a diverse phenomenon, including "spoken" sentences and unworded speech (Hurlburt and Akhter 2006).

Like other conscious experiences, there is a sense of ownership or "mineness" for episodes of conscious thinking like inner speech. As Joelle Proust puts it, when one is thinking, "there is a sense of being the thinker of that thought, and a sense of this thought as being the one that is presently occupying one's attention" (2009: 253).

What about the sense of agency? When it comes to our bodily actions, it is easy to motivate the distinction between ownership and agency. Think of the difference between an intentional action, such as walking across a room, and nonintentional bodily movement, such as being pushed. In both cases one will experience the movement as one's own, but only in the former will it be experienced as one's own doing. But things are not so simple when it comes to psychological ownership. No such contrast comes readily to mind. What would it mean to have one's thoughts *done* to one, rather than being produced by oneself? If we are to get the idea of a sense of agency into view by contrasting cases in which it is present and absent from our own experience, then it seems we are unable to get a grip on it. Furthermore, many of the concepts of intentional action, like planning, choice, practical reasoning, and deliberation, get no purchase for conscious thinking, except perhaps in fairly *recherche* cases. As Christopher Peacocke puts it, "...when you think a particular thought" (1999: 209). Intentions in advance aren't required for every intentional action, of course, but the absence of intentions in advance contributes to the difficulty in drawing a contrast between agency and ownership. Quite generally, conscious thinking feels unplanned and spontaneous. How could it be otherwise, since planning is itself a form of thinking?

This might suggest that in the psychological domain only ownership matters. However, it is worth taking seriously the idea of a sense of agency for conscious thinking, for two reasons. First, it is part of our ordinary understanding of the mental

<sup>2</sup> For an opposing view, see Macherly (2018), who argues that because inner speech is communication, and so governed by pragmatic norms, it isn't an instance of thinking.

that thinking is something we do rather than something that happens to us. Even granting the respects in which it differs from bodily action, there is a vivid, if vague, sense that conscious thinking is a mental action. Not only do we intuitively *believe* that conscious thinking is an action, but we *experience* it as such. When I think about how to word a sentence or structure a section of a paper, I experience the activity as one over which I have discretion and control, even if it isn't the sort of thing about which I might plan or form intentions in advance.<sup>3</sup>

Second, the phenomenon of delusions of inserted thoughts, experienced by some schizophrenic patients, has suggested to some that the distinction between ownership and agency arises in the same way for thinking as it does for bodily action (Campbell 1999; Proust 2009). Consider this well-known example:

I look out the window and I think that the garden looks nice and the grass looks cool, but 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)

These subjects report thinking thoughts that are not their own. The thoughts are experienced as their own, in the sense that they occur as events within their stream of consciousness. But these patients deny that they are the source of their own thoughts.<sup>4</sup> Instead, they claim, another's thoughts, often with disturbing content (e.g. "Kill God!"), intrude on their thinking. Philosophers and psychologists have argued that we should try to understand the experience of these patients in a way that renders these reports intelligible. *Prima facie*, that involves understanding how it could so much as *seem* to one that an act of thinking one experiences is not one's deed. The natural thought is that whatever underlies the sense of agency for conscious thinking is absent in these cases. The delusion of inserted thoughts is grounded on an illusion of the absence of agency for one's own thinking. As John Campbell puts it: "At the very least, these reports by patients show that there is some structure in our ordinary notion of the ownership of a thought which we might not otherwise have suspected" (1999, 610).

There is good reason, from both psychopathology and folk psychology, to try to understand the sense of agency for conscious thinking. In the ordinary case, we attribute our acts of conscious thinking to our active natures. Our question is the following: what grounds these self-attributions of agency? What justifies us, subjectively, in taking particular acts of conscious thinking as our deeds? Notice that this question focuses on particular cases. The question is not: why do we conceive of conscious thinking, quite generally, as an action? It is rather: why, when I hear the

<sup>3</sup> This is consistent with denying, as I did in the last paragraph, that we can get a grip on the idea of a sense of agency for conscious thinking through reflection on phenomenal contrast cases, as we might for intentional bodily action.

<sup>4</sup> See Saks (2007) for a compelling memoir about schizophrenia and Mellor (1970) for discussion of its symptoms.

sentence, “this sentence needs a compelling example”, uttered in inner speech, am I justified in taking that as my own deed?

The paper proceeds as follows. After identifying a few key assumptions, the next section motivates the following constraint on an account of the sense of agency:

**PERVASIVENESS:** Whatever underlies the sense of agency is a pervasive feature of our conscious cognitive lives.

Section 3 sets out what I call the “agency as endorsement” model for explaining the sense of agency and argues that it cannot satisfy Pervasiveness. In Sect. 4 I sketch an account of the epistemology of conscious thinking and draw on it to propose a novel explanation of the sense of agency. Section 5 explains how the account satisfies Pervasiveness.

## 2 Pervasiveness

It is difficult to say anything uncontroversial about the senses of agency or ownership. I will mention a few assumptions that guide the following. First, the phenomenon is that of “taking” some act as one’s doing. This formulation is deliberately neutral, leaving it open whether the taking is doxastic, intuitional, dispositional, or something else. Second, as mentioned, the methodological approach to understanding the sense of agency pursued here is to answer the following personal-level, normative question: what grounds our self-attributions of agency to acts of conscious thinking? (Bermudez 2019) In virtue of what are we subjectively justified in claiming that a stretch of thinking is our own doing?<sup>5</sup> This contrasts with a causal, sub-personal question: what mechanisms underlie the sense of agency? (cf. Frith 1992; Campbell 1999). This question is pressing, too, of course. But I will focus on the normative question.

I have formulated both assumptions in a way that aims to resist the idea that agency is somehow written into the face value of an experience of thinking or is represented in the content that constitutes its phenomenal character. This is because I assume a deflationary or reductive account of the senses of agency and ownership. This contrasts with inflationary views, which hold that the senses of agency and ownership are realized as a distinct aspects of the phenomenology of the relevant conscious experiences (Zahavi and Kriegel 2015). On this view self-reference is built into the *character* of our conscious experiences. Thus, awareness of a mental act as my doing requires introspecting its character, recognizing the quale of agency, and taking it at face value. By contrast, the reductive view denies that mineness or agency are distinctive features of the phenomenology of the relevant experiences. I won’t argue against the inflationary view here (see Bermudez 2011, 2019; Schear

<sup>5</sup> That leaves open the question whether we are in fact justified in taking our conscious thinking as our own doing. Perhaps we are under an illusion. I assume we are not, and am inclined to assume that an account of the sense of agency ought to explain our objective justification for taking thinking as our deed. But I won’t pursue that here.

2009). But notice that those who have denied the existence of a distinctive quale of ownership or agency (e.g. Bortolotti 2009) can accept the idea that there is a sense of agency for conscious thinking, as I understand it. Furthermore, one can reject the inflationary approach without denying that reflection on conscious experience is relevant to understanding the sense of agency.

Those are the three assumptions that guide the following. I will now motivate a constraint on an account of the sense of agency for conscious thinking. The constraint is latent in much of the literature and should be acceptable to many.

Much of the literature on the sense of agency for conscious thinking accepts the following:

CAMPBELL'S THESIS: In delusions of thought insertion, the hypothesis of thought insertion is explained by the absence of the sense of agency for one's own thinking.<sup>6</sup> (Campbell 1999; Hoerl 2001; Bortolotti 2009; Proust 2009; Pickard 2010; Fernandez 2013; Bermudez 2019)

Subjects who experience delusions of thought insertion deny that they produce some of the thoughts they experience. The hypothesis motivating much contemporary work on the sense of agency is that these subjects make these reports because they do not experience the sense of agency for these acts of thinking. Whatever grounds the sense of agency, it is absent in these cases. The thesis is restricted to explaining the *hypothesis* of inserted thoughts because the absence of the sense of agency alone cannot be the whole story in explaining the delusions. That's for two reasons. First, the absence of the sense of agency cannot explain why these subjects maintain their delusions in the face of all the available counterevidence.<sup>7</sup> That it *seems* to one that one is not the agent of one's own thinking might explain why it occurs to one that one is not. But it cannot explain why one *persists* in holding onto the belief that one is not.<sup>8</sup> Second, the absence of the sense of agency cannot explain the specific content of the delusion. It can explain why the patient doesn't attribute the thought to her own agency, but it cannot explain why she attributes it to some specific other person's agency (like Eamonn Andrews, in the example above.) As we might put it, it can explain the *negative* content of the delusion (that it is not my deed) but not its *positive* content (that it is so-and-so's deed).<sup>9</sup> Campbell's thesis, then, claims that the absence of the sense of agency explains why it seems to some schizophrenic subjects that a thought is not the product of their agency.

<sup>6</sup> So-named for John Campbell, whose 1998 argued for the thesis and influenced much of the contemporary work that tries to answer the normative question about the sense of agency by a backwards inference from reflection on the inserted thought delusion.

<sup>7</sup> This is a point made familiar by so-called "two-factor" accounts of delusions. See Davies et al. (2001).

<sup>8</sup> This assumes delusions are beliefs, of course. But nothing here hangs on that assumption. See Bortolotti (2009) for discussion.

<sup>9</sup> I remain neutral here on the explanation of the positive content of delusions of thought insertion. Karl Jaspers notoriously claimed that delusions are "ununderstandable" (1913). Much of the literature on delusions has denied this, and aimed to make sense of how delusions arise (what Jaspers calls achieving "genetic understanding" or empathy, *ibid.* 17) (see Thornton 2007 for discussion.) But one thing that Jaspers might be taken to mean is that the specific positive content of delusions cannot be given an empathetic, rationalizing explanation. (I'm indebted to a conversation with Matt Parrott for this.)

Campbell's thesis, and some uncontroversial facts about delusions of thought insertion, give us reason to accept the following:

**PERVASIVENESS:** Whatever grounds the sense of agency is a pervasive feature of our conscious cognitive lives.

Delusions of thought insertion are extremely rare; non-schizophrenic subjects don't experience them, nor are they inclined toward them. (Remember: we can't grasp the idea of a sense of agency for conscious thinking through reflection on cases where it is absent.) If the absence of the sense of agency for conscious thinking explains delusions of thought insertion, then, given that they are very rare, it must be very rare for the sense of agency to be absent. And so the sense of agency must be pervasive.

By a "pervasive feature of our conscious cognitive lives" I mean that whatever underlies our attributions of agency to thinking is found more or less whenever we are engaged in conscious thinking. Again, that should not be taken to imply that there is a distinctive *qualé* that accompanies any act of thinking, such as the feeling of trying or exerting oneself (though such a feeling may attend some acts of thinking.) There needn't be something it is like, all the time, to be the agent of one's thoughts in order for there to be a feature of our conscious thinking, quite generally, in virtue of which it makes sense that we attribute our acts of thinking to our agency. The claim is that whatever explains why we take thinking to be something we do, it isn't something occasional or fleeting, but rather something that always, or nearly always, accompanies thinking.

Campbell's Thesis gives us reason to accept Pervasiveness. One may hold that the absence of the sense of agency is only part of the story here, accepting Campbell's Thesis but rejecting Pervasiveness. I think this is against the spirit of Campbell's Thesis, but no matter. One can accept Pervasiveness while rejecting Campbell's Thesis. Perhaps absence of the sense of agency doesn't explain delusions of thought insertion. Perhaps these delusions should be given an entirely different kind of explanation (Roessler 2013). Pervasiveness is independently plausible. Again, there is no clear analogue of the intuitive contrast between active and passive bodily movements that motivates the claim that there is a sense of agency for bodily action. We can draw on something like Wittgenstein's question (what is the difference between my arm going up and my intentionally raising my arm?) to motivate the claim that there is a sense of agency for bodily action. But there isn't such an intuitive contrast for conscious thinking. Earlier I mentioned this in order to flout the possibility that there isn't a sense of agency for conscious thinking. But that is wrong. If there is such a sense, and we cannot get a grip on the contrast, that would suggest that agency is pervasive. It is the absence of agency for thinking on which we cannot get a grip.

This is less controversial than it may appear. It is undeniable that conscious thinking needn't be *planned* or *deliberate*. It needn't be something we self-consciously and attentively pursue and exercise control over throughout its duration. Conscious direction isn't required for conscious thinking or imagining. Still, two points are worth emphasizing here. First, even when we daydream or let our thinking stray, the activity is under our control in that we can terminate it if we are interrupted. Daydreaming is like tapping one's fingers or rolling one's tongue over in one's mouth:

a kind of undirected activity that is within one's active control even if it isn't initiated through conscious attention and practical reasoning. Second, even if one is surprised, even alarmed by one's fantasies, we have a sense that they are things that we are up to, something that can be attributed to us and something for which we are responsible, in at least the sense that we might be the fitting objects of some sort of condemnation or praise for them. I will have more to say about purported cases of passive thoughts in part five below. My point here is simply that one can accept Pervasiveness while rejecting Campbell's Thesis.

### 3 Agency as Endorsement

Pervasiveness is intuitively plausible and supported by Campbell's Thesis, which is widely accepted in the literature. So Pervasiveness ought to be acceptable to many who have proposed answers to our question. Yet at least one prominent model is incompatible with it. That gives us reason to reject the model and reveals that Pervasiveness is a constraint with some teeth.

I call the model "agency as endorsement"; versions have been presented by different philosophers (Graham and Stephens 2000; Hoerl 2001; Bortolotti 2009; Fernandez 2010; Pickard 2010; Bermudez 2019). Here is the basic idea. Taking a thought to be one's own deed consists in taking it to possess some normative status, such as being intelligible, justified, or otherwise reflective of one's point of view. This means that the act of taking the thought as one's own is realized in a higher-order reflection on the *content* of that thought. Different philosophers understand this higher-order reflection in different ways. Graham and Stephens draw on Frankfurt's (1971) idea of a second-order desire, whereas Hoerl, Bortolotti, Fernandez, and Pickard draw on Moran's (2001) account of self-knowledge. Despite these differences, there is a core view that all share. Here are two representative passages:

... an important part of what it is to take an occurrent thought as one's own is to take it as something that one can deploy in conscious reasoning. Ownership needs to be understood in terms of the active process of thinking.

Bermudez (2019, 277)

... there is a particular experience that one will have if one determines that one has a belief from the first-person perspective—namely, one will experience that belief as matching the world... That belief is presented to me as being correct.

Fernandez (2013, 168)

The core idea, then, is that taking an act of thinking as one's deed involves taking it as an expression of an *attitude* which one takes to have some *status*, such as being true, responsive to my reasons, or suitable for deployment in reasoning.<sup>10</sup> Defenders of this view tend to focus on conscious judgments, so I will do the same.

<sup>10</sup> Thanks to an anonymous referee for help clarifying how best to capture the view.

The agency as endorsement model is in line with the assumptions listed earlier. The approach is deflationary because it doesn't assume there is some distinctive feeling of agency or ownership. Instead, the content of a given act of thinking is such that it strikes one as one's own. This requires only that one be able, on reflection, to recognize it as, for example, in line with one's point of view. The model also presents an answer to the normative question. One is justified in taking a conscious thinking as one's own deed because one finds the content of thought intelligible or in line with one's point of view. Furthermore, the view is neutral on what kind of act endorsement is. It might be a higher-order judgment, a disposition to judge, or something akin to an intuition. Finally, the model is thought to explain delusions of thought insertion. Here are some representative passages:

Part of the cognitive dissonance experienced by schizophrenic patients with these symptoms is that they experience (and are able to ascribe to themselves) thoughts that they do not know how to integrate with the rest of their cognitive and affective lives. A natural explanation of this cognitive dissonance... is that only occurrent thoughts that can feature as premises in reasoning are taken to be self-originated.

Bermudez (2019, 277)

I propose that schizophrenics disown mental events that seem to be manifestations of mental states that they do not, for some reason or other, endorse. Taking up the practical stance, looking outwards to the world, they judge that the mental states which these thoughts, impulses, or feelings bring to consciousness are not warranted or appropriate: they do not reflect how the world actually is or should be.

Pickard (2010, 67)

The proposal is now that this experience [of a belief's being true], which normally accompanies self-knowledge, is missing in thought insertion. When the patient is aware of those beliefs that she disowns, she does not experience that those beliefs are forcing her to accept any particular picture of the world. That is, she does not feel compelled to endorse their contents. This is why the patient affirms that those beliefs are not hers.

Fernandez (2013, 167–8)

A person denies that she is the agent of a given thought because she finds that she cannot explain its occurrence in terms of her theory or conception of her intentional psychology.

Graham and Stephens (2000, 162)

Again, there are differences between the accounts. Bermudez appeals to taking a thought as a premise in reasoning, Pickard talks of endorsement, Fernandez refers to an experience of a belief's seeming to be true, and Graham and Stephens speak of a thought's fit with one's understanding of one's own psychology. There may be cases where these come apart. But it seems fair to assume that they are extensionally very close notions. One is disposed to use as premises in reasoning propositions one would, on reflection, endorse. If the thought is the expression of a belief, then endorsement involves taking the thought to be true. And presumably one's understanding of one's own intentional psychology is informed by what attitude

one would find true, on reflection. Furthermore, the basic idea in play, that taking a thought as one's deed is a matter of reflectively taking it to possess some normative status or other, is common ground.

Explaining the sense of agency involves identifying a feature of our conscious thinking that would justify us in taking our thoughts to be our deeds. If Campbell's Thesis is true, then that feature must plausibly be absent from cases of inserted thoughts. And by Pervasiveness, it must be a feature of all or nearly all of non-schizophrenic acts of thinking. The agency as endorsement model has considerable plausibility in explaining delusions of thought insertion. As mentioned, the thoughts that are reported as inserted typically have disturbing and bizarre contents, presumably thoughts that are out of keeping with the subject's background beliefs and point of view. It is plausible, then, that cases of delusional thoughts are cases where a subject would not endorse the content of the thought. So it looks like the model is in line with Campbell's Thesis. However, it is not consistent with Pervasiveness.

I have many spontaneous thoughts the content of which I would not on reflection endorse as justified, true, representative of my broader point of view, which I would not deploy in reasoning, and yet which I recognize as fully my own. I assume I am not unusual in this respect.

To avoid autobiography, let's talk about our mutual friend, Dianne. Dianne has a dreadful co-worker with an office down the hall, Rebecca. In order to get to lunch, Dianne needs to pass Rebecca's office. But if Rebecca sees her, she will invite herself to lunch, ruining the whole thing. On the other hand, if Dianne stays in her office, she will get hungry and bored, unable to concentrate on her work. To help convince herself to risk passing Rebecca, Dianne engages in a little fantasy narrative about all the terrible things she could do to avoid eating lunch with Rebecca. In the midst of this, she thinks, "The thing to do is simply push her out the window!" Now, we may safely assume that Dianne does not take it to be *true* that this is the thing to do, nor would she be inclined to use this as a premise in practical or theoretical reasoning. Yet there is nothing alien about this thought. Though she might be put off by the violence of her thought, Dianne will recognize her own agency lying behind it.

Bermudez, Pickard, and Fernandez' proposals seem extensionally adequate. It is easy enough to see why. Their views work well when we think about conscious thoughts that express or manifest a subject's beliefs. (And even here it is plausible that there are mundane cases in which one recognizes a thought as the manifestation of belief but refrains from endorsing it. (Leite 2018)). But many of our acts of conscious thinking are not expressions of belief: they involve imagination, fantasy, and the like. Much of our conscious thinking is not truth-directed and does not aim to represent our considered stance on the matter at hand. Often we simply think things to ourselves for our own entertainment. If agency is a pervasive feature of our conscious thinking, then endorsement looks like the wrong model for agency, since endorsement is not pervasive.

One might worry about the example.<sup>11</sup> Perhaps Dianne's act of thinking is the expression of pretence or pretend belief, rather than judgment. If that is correct, then

---

<sup>11</sup> Thanks to an anonymous referee for this.

we may suppose that she does endorse the thought as true in the fictional world of her pretense, if pretense is a cognitive state much like belief (Nichols and Stich 2001). Pushing Dianne out the window isn't the thing to do simpliciter, of course, but perhaps it is the thing to do, under the suppositions that constrain Dianne's pretense.

However, the sense of agency for acts of conscious thinking doesn't require knowledge of the attitudes that thinking manifests or expresses.<sup>12</sup> Suppose that Dianne loses track of what she supposes and what she believes, and takes her conscious thought as the expression of a judgment. Still, I contend, she will recognize it as a manifestation of her agency. Or, perhaps more plausibly, suppose that Dianne withholds belief about whether she believes that the thing to do is push Rebecca out the window. That is, suppose that she is unsure whether that thought reflects what she believes or is merely the product of pretense. It is doubtful that recognizing the thinking as her doing awaits her settling this question. Indeed, it seems fair to suppose that the question of what attitude is expressed by her act of thinking presupposes that the thinking is her deed.

Graham and Stephens' account might be thought to fare better in this respect. On their view a subject recognizes herself as the agent of a thought when that thought conforms to her understanding of her own psychology. And this could be true of Dianne. She doesn't think it is true that she should push Rebecca out the window; it does not "reflect how the world actually is or should be", as Pickard puts it. But thinking that thought is plausibly an expression of her desire to avoid eating with Rebecca, as well as her anger and annoyance at her co-worker. So if we interpret endorsement in a more self-directed way, as intelligible in light of my understanding of my psychology, then our case isn't a counterexample.

Again, the difficulty is that Graham and Stephens' account will only work in cases where a subject has a more or less *accurate* conception of their own psychology. But it seems unproblematic to imagine a case where a subject lacks the knowledge of the mental states that render her act of conscious thinking intelligible, but treats that thinking as manifesting her own agency. Suppose that Dianne is self-deceived about her feelings toward Rebecca and her desires to avoid her. Dianne likes to think of herself as a perfect co-worker and loving toward everyone. Planning her lunch break, she imagines walking down the hall, past Rebecca and suddenly thinks to herself: "I ought to push her out the window!" This is an expression of her feelings and desires, but Dianne is not in a position to appreciate this. She is ignorant of these feelings and desires. Graham and Stephens' view predicts that Dianne will deny that this thought is a manifestation of her agency. However, though it might seem strange to her, I doubt that she would view things this way. She would take it, instead, that her imagination took a somewhat unexpected turn.

The point here doesn't depend on our intuitions about particular stories. All versions of the agency as endorsement model are extensionally inadequate. Many of our conscious thinkings have contents that we would not endorse, take as true, or

<sup>12</sup> Furthermore, if you think that we come to know our attitudes by introspecting inner speech (Machery 2018), then it looks like the agency as endorsement model has things backwards.

use as premises in reasoning. And many thoughts are motivated by mental states about which we are ignorant. We are not omniscient about the mental states that motivate our conscious thoughts, and there is no single normative status (truth, justification, etc.) which we would attach to all of our acts of thinking. Yet the sense of agency is a pervasive feature of our cognitive lives. So taking a conscious thought as a manifestation of agency cannot be a matter of endorsing it.

#### 4 An Epistemology for Conscious Thinking

I think that Bermudez is right about the cognitive dissonance experienced by schizophrenic patients. As he puts it: "...they experience (and are able to ascribe to themselves) thoughts that they do not know how to integrate with the rest of their cognitive and affective lives." (ibid.) But Bermudez errs in explaining this dissonance by appeal to the *content* of punctate thoughts in the stream of consciousness and higher-order reflection on them. There are other ways of integrating acts of thinking with the rest of one's cognitive life. This section proposes an alternative. The basic idea is that a thought is recognizable as one's deed because it is recognizable as a part of an unfolding extended cognitive activity. The suggestion is that in cases of inserted thoughts a subject cannot recognize an act of thinking as a part of something that she is up to. It is a disruption of her own cognitive activity, not a disruption into what she takes to be true or in conformity with her self-conception.

The account draws on an epistemology for conscious thinking which I have developed elsewhere (Doyle 2018). Here is a sketch.

It is a striking feature of work on the sense of agency for conscious thinking that it restricts its focus to punctate acts of thinking, the occurrence of single sentences or small thought tokens in consciousness. To be fair, this is encouraged by the cases of thought insertion many set out to explain; for these are often reported as short bursts of thinking that disrupt the subject's experience. Still, attending only to punctate thought occurrences is problematic. We get a single act of thinking into view and ask, "where is the agency in it?" Given that it is implausible to suppose that each act of conscious thinking is itself the product of practical reasoning, it is natural to suppose that agency is realized in the subject's relation to the thought's content rather than its production.

Gilbert Ryle complained about restricting our focus to punctate acts of thinking in reflecting upon the nature of mental activity. He writes:

It is often supposed by philosophers and psychologists that thinking is saying things to oneself, so that what *le Penseur* is doing on his rock is saying things to himself... Very likely *le Penseur* was just murmuring something under his breath or saying it in his head... The description "murmuring syllables under his breath", though true, is the thinnest possible description of what he was engaged in. The important question is "But what is the correct and thickest possible description of what *le Penseur* was trying for in murmuring those syllables?"

Ryle (2009: 501)

Now, “correct” trivializes Ryle’s point. And it isn’t clear why we should seek the thickest *possible* description. But Ryle is clearly onto something. Conscious thinking typically takes the form of a series of discrete episodes that constitute a complex cognitive activity. We must distinguish the discrete episodes of thinking in which a particular content is represented in phenomenal consciousness and is present “before the mind’s eye” from the extended activities of which these episodes form a part. In the typical case, a thought that is expressed in inner speech is part of an ongoing cognitive activity. You say something to yourself, *in sotto voce*, while in the midst of, e.g. answering a question, jogging your memory, reciting a recipe, outlining a paper, distracting yourself from an annoying person, and so on.

This distinction is important for the epistemology of thinking because one cannot know of extended cognitive activities by introspecting the character of one’s inner speech. Suppose that I am answering a question about inserted thoughts. As I try to come up with my answer, I utter various thoughts in inner speech. Let’s suppose further that one of the things I mutter to myself is: “We should try to understand what schizophrenic subjects experience is like that they report inserted thoughts.”<sup>13</sup> Now, if you ask me: “what are you thinking?” I can answer in at least two ways. I might report on the specific thought that was running before my mind’s eye, perhaps by introspecting my conscious experience, eavesdropping on silent soliloquy, as Ryle put it (Ryle 1949: 148). Or I might cite the extended activity of answering a question. Introspection can reveal the content of my thought, but it cannot reveal that I thought it in the midst of answering a question about inserted thoughts. This is what we should expect if we take the analogy between inner and outer speech seriously. If you heard me utter that sentence, and lacked the relevant background knowledge, you would not thereby be in a position to know that I was answering a question, imagining an objection, reciting a poem, or anything else.

In the epistemology of conscious thinking we need to distinguish punctate episodes of inner speech from what I have called extended cognitive activities. When we know the latter, it is not by introspecting our experience. Instead, I propose, our knowledge of extended cognitive activities is practical knowledge or knowledge in intention. Practical knowledge is noninferential and nonobservational knowledge of one’s own intentional actions (Anscombe 1957; Velleman 1989; Moran 2004; Setiya 2007; O’Brien 2007). However, these are only negative characterizations of practical knowledge. I remain neutral on many controversial issues. As I understand it, practical knowledge is “knowledge in intention” in the very weak sense that intending to do A can itself provide a subject with an entitlement to self-ascribe the action of doing A, now or in the future. This is consistent with the possibility that other evidence may defeat the entitlement provided by the intention, such as evidence that one is unlikely to succeed. And it leaves room for different ways of understanding the entitlement and more besides. For example, some philosophers, taking the

<sup>13</sup> It is rare, of course, for one’s inner speech to be so elaborate. Fernyhough (2016) points out that inner speech is often compressed, akin to personal notes. But presumably anything one can say in outer speech one can mutter in inner speech. So there is no harm imagining a case like this, even if it is atypical in certain respects.

phrase “knowledge in intention” literally, hold that an intention can itself constitute knowledge either because intention is constituted by belief and subject to the standards of theoretical rationality (Velleman 1989) or because practical knowledge is a *sui generis* form of knowledge subject only to the standards of practical rationality (Anscombe 1957). This contrasts with the view that intending to do A or doing A intentionally can themselves be reasons for the self-ascription of action, though one can intend to do A without possessing practical knowledge (O’Brien 2007). The proposal, then, is that knowledge of extended cognitive activities (e.g. “I’m answering a question.” “I’m trying to remember when the restaurant opens.” “I’m imagining how an interlocutor might object to my argument.”) is nonobservational and noninferential knowledge in intention.

Earlier I considered reasons for being sceptical about a sense of agency for conscious thinking. But that scepticism was targeted at agency for punctate acts of inner speech. The idea that extended cognitive activities are intentional actions should be uncontroversial. Extended cognitive activities have many of the marks of intentional actions. It is appropriate to ask the subject engaged in such an activity why she is so engaged, and her answer will give her reason for action (Anscombe 1957). These activities can be given rationalizing explanations which cite the beliefs and desires of the subject (Davidson 1980). They are goal-directed activities, and the goals are those determined by the agent herself (Soteriou 2013). Finally, we have nonobservational and noninferential knowledge of these activities. When asked, one can straightaway report that one is jogging one’s memory or reciting *Ozymandias*, or what have you.

Fair enough, one might think. But what does this have to do with the sense of agency? Even if the epistemology of conscious thinking should operate with an expanded diet of examples, questions about the sense of agency for conscious thinking are questions about punctate acts of thinking, paradigmatically inner speech, the sorts of events reported by schizophrenic patients. How does the epistemology of extended cognitive activities bear on that?

Start with an analogy. Suppose that you are walking to a nearby riverbank. Walking to the river is an intentional action and you have practical knowledge of this action. Without observation or inference you could straightaway tell me what you are up to. As you walk, your feet and limbs move about in various subtle ways to keep your balance, find your footing, and so on. Take one fine-grained description of the moment of your right foot as it gains purchase on a rock. This is a motion done in the service of satisfying your intention to walk to the river. It is part of the extended activity of walking to the river. And it would be perfectly appropriate for you to say, of the movement, that you did it intentionally, or that you meant to do it. Despite this, it would be wrong, I think, to say that you had a further intention which that particular motion satisfied (or failed to satisfy.) You are moving the foot intentionally because you intend to walk to the river, not because you intend that you move your foot in that particular way.<sup>14</sup> It helps here to keep in mind the idea that an intention sets a standard for performance (Anscombe 1957; Soteriou 2013). It

<sup>14</sup> See McDowell (2011) and Bratman’s rejection of the “simple view” in his 1987.

should be clear that, at least in the ordinary case, the only standard that bears on the movement of the foot is whether it aids in the achievement of what you intend to do: walk to the river. But in order to do *that* it needn't move in any one particular way, we may suppose. Were the foot to move in a slightly different way, we wouldn't have grounds for claiming there was a failure in performance, unless the subtle change prevented you from walking to the river. So it is plausible to suppose that there are parts of intentional activities that are themselves not intentional, in the sense that they are the executions of intentions or sub-intentions. If that is correct, then knowledge of fine-grained parts of intentional actions cannot be knowledge in intention. So we might ask: how are they recognizable to us as our deeds if our awareness of them is not an instance of practical knowledge?

They are recognizable to us as our deeds because they are recognizable as parts of the unfolding action (Gibbons 2010; Schwenkler 2011). When you perceive your body moving a particular way, you are not simply observing it, but you also don't know the movement in virtue of intending to make that specific movement. Rather, your intention to walk to the river enables you to have a practical perceptual awareness of the movement. It is immediately perceptually recognizable to you as your action and as in the service of your intention. Just as background knowledge might enable you to possess noninferential knowledge by perceiving, say, a brand of car or a foul ball, so too can your intention enable you to perceive a particular movement as your own action, even if you do not have an intention to perform that particular movement (Gibbons 2010).

Things work exactly the same way with conscious thinking. You ask me a tough question about inserted thoughts. As I ruminate, I think, "We should try to understand what schizophrenic subjects' experience is like that they report inserted thoughts." I didn't set out to think that specific sentence, in the sense that I did a bit of syllogizing that led to it. I really do eavesdrop on silent soliloquy. But it isn't exactly eavesdropping, since it isn't like listening to someone else speak. We should distinguish inner speaking from inner hearing (FERNYHOUGH 2016), but inner hearing is practical. That's because introspecting my own thinking is informed by the intention that governs the extended activity. I know what I am up to and so I am aware of the act of thinking as a part of that extended activity. It seems to me that I am doing something when I hear myself intone those words *in sotto voce* because I already know, or am in a position to know, that I am up to something: answering a question. We recognize acts of thinking as our own deeds because we can integrate them with our practical understanding of our own ongoing activities, activities that extend beyond the punctate thoughts manifest in consciousness at a time.

## 5 The Sense of Agency

The proposal, then, is that the sense of agency for acts of conscious thinking is underwritten by intentions to engage in extended cognitive activities. Acts of conscious thinking seem like things you do because they seem like parts of activities that you are engaged in. Again, Bermudez is right to stress that the sense of agency has to do with integrating acts of thinking with our cognitive lives. But he is wrong

to suppose that this requires integrating punctate attitudes into our conception of the true and the good and the like. Instead, we are aware of thoughts as our deeds insofar as we can integrate them into our understanding of our own unfolding actions, actions of which punctate thoughts are a part. To do this we needn't endorse them or know what attitudes are expressed by them. We can be wrong about our attitudes. All that is required is that the thought be recognizable as an element in a cognitive activity, including those, like imagining, that are not truth-directed at all.

This proposal is in line with the assumptions mentioned earlier. The account is deflationary because we explain the sense of agency by appeal to a form of self-awareness made possible by intentions, rather than a quale of agency. It is neutral about how exactly to understand taking one's own thinking as a deed. And it offers an answer to the personal-level question about how the subject is entitled to take a particular act of conscious thinking as a deed.

Although it is beyond the scope of this paper, and requires work that cannot be done from the armchair, I think the proposal is a promising way to think about delusions of thought insertion. The basic idea here is that these subjects experience disintegration of their own ongoing conscious activities. The reported thoughts are seen as alien because they disrupt cognitive activities underway and aren't recognizably a part of what they are doing. The thoughts are attributed to another because they do not answer to the question, "why are you thinking that?" where the answer would explain what one was up to in thinking it. Vindicating this would require better understanding the relation between thought insertion delusions and other disruptions of agency in schizophrenia.

Furthermore, the proposal has an easy time with the counterexamples to the agency as endorsement model. Think back to Dianne. She thinks to herself that she ought to push Rebecca out the window. She doesn't endorse that thought as true or as a suitable ground for inference. If she is self-deceived, she won't see it as conforming to her self-conception. But she will be able to recognize it as part of one of her ongoing activities: imagining some pleasing hypotheticals for dealing with an annoying co-worker.

Of course, this all depends on the claim that what I have called extended cognitive activities are a pervasive feature of our cognitive mental lives. One might doubt this. Is intentional action really a pervasive feature of mental activity? The two most obvious purported counterexamples are isolated thoughts and activities that are not intentional. I will consider them in turn, then I will briefly consider one further worry.<sup>15</sup>

---

<sup>15</sup> Other purported counterexamples are obsessive thoughts and mind-wandering. See O'Brien (2013) for a compelling defense of the claim that obsessive thoughts are active phenomena. See Metzinger (2013) for the view that mind-wandering involves a loss of autonomy and Irving (2016) for the view that it is unguided thinking. Finally, see Shepherd (2019) for the claim that mind-wandering is a search for a more rewarding goal guided by the cognitive control system. Although his view is more amenable to the present account, Shepherd notes that in the relevant cases subjects do not report an intention to engage in mind-wandering. One question to consider here is whether we possess practical knowledge of the constituent elements of mind-wandering. One may not intend to wander, but one may have practical knowledge of recalling this or and that, and then imagining something, and so on. One's action isn't intentional under the description "mind-wandering", but its constituent elements are. (See my discussion of O'Shaughnessy on daydreaming for this option.) Unfortunately, I don't have the space to consider this here. Thanks to an anonymous referee for raising this issue.

The idea of a thought that just “comes to one” is common, but a thought that is wholly isolated from one’s ongoing intentional activities is rarer than this suggests. This is easier to see once we recognize that the line between mental and bodily actions is often blurry; many of the activities of which a conscious thought is a part are realized in thought and action. It’s obvious that we can realize our cognitive actions, in part, physically, such as when you count with your fingers. But mental actions can also be a part of physical actions. Thinking, “I’m hot!” might be a part of noticing yourself sweating, distracting yourself from an unpleasant conversation, ensuring one’s guests are comfortable, and so on. What matters here is not the realization of a series, but the proper application of a “thick” description that picks out an extended intentional action.

Ryle claims that thought is, as such, “incipient” (2009, 416). “Roughly, a thought comprises what it is incipiently, namely what it is the natural vanguard of. Its burden embodies its natural or easy sequel.” (ibid.).<sup>16</sup> Even seemingly stray thoughts are caught up in extended activities, which we might need to stop and reflect on to make explicit. Here is one of his examples:

... I ask a tired visitor from London what he has been thinking about. He says, ‘Just about the extraordinary peacefulness of your garden.’ If asked, ‘Than what do you find it so much more peaceful?’ he replies, ‘Oh, London, of course.’ So in a way he was thinking not only of my garden but of London, though he would not, without special prompting, have said for himself that he had had London in mind at all.

Ryle (2009, 413–414)

Again, to say that thought is as such incipient, or, as I have put it, that extended cognitive activities are pervasive, does not require that a complex series of thoughts is realized every time we think. It requires only the applicability of a thicker description of an act of thinking. When struck by the peacefulness of the garden, the Londoner experiences what might seem a stray thought. But it is part of his activity of comparing the garden with London (which may have unfolded in his inner speech) as well as the activity of enjoying the garden (which needn’t have involved further acts of inner speech).

Daydreaming is a cognitive activity, but one might doubt that it is an intentional cognitive activity of which we possess nonobservational knowledge. Suppose we focus on a case of daydreaming that involves inner speech. There is a general reason for supposing that this too must be active in the relevant sense. It is plausible to conceive of inner speech, in the Vygotskian image, as the interiorization of outer speech. But then it will also seem natural to claim, with O’Shaughnessy, that conscious thinking is “a case of ‘talking to oneself’, and being talking cannot but be intentionally active” (2000, 217). And even if one is reluctant to simply draw that inference, the connection between inner and outer speech will help to make clear why daydreaming is also an extended cognitive activity. As O’Shaughnessy points

<sup>16</sup> Or, as Brian O’Shaughnessy puts it: “...the main processive constituents of the stream of consciousness of the conscious are intentionally active phenomena” (2000, 200).

out, as with outer speech, a subject who engages in conscious thinking is in a position to access “the rationale from moment to moment of the progression” of her thinking (ibid 218-9). I take it his point is that, though one might not intentionally daydream, one has nonobservational practical knowledge of one’s action under a different description, one that specifies the connections between the various things that run through one’s mind. O’Shaughnessy puts it thus:

... the mind remains intentionally active under other headings at each point in that process. Thus, the immediate future is already prefigured in descriptions like ‘I was remembering the time we visited X...’, and this strongly suggests the existence of a foreknowledge conferred by intentions.

Ibid. 219

Even though you don’t set out to daydream, you can know, straightaway and without observation, what you are doing. An attractive explanation of this epistemological feature of daydreaming is that it is an intentional phenomenon.<sup>17</sup> Both the nature of conscious thinking, paradigmatically in the form of inner speech, and its epistemology give us reason to believe that it is a pervasively agential phenomenon.

One final worry is worth addressing.<sup>18</sup> We are not infallible and so sometimes lack self-knowledge. Some versions of the agency as endorsement model require self-knowledge of our attitudes. Taking an act of thinking as one’s deed requires recognizing it as the manifestation of an attitude one endorses or as consistent with one’s theory of one’s psychology. This predicts that in cases where one is in error or ignorant, one will not take the act of thinking as one’s deed. Since this can happen, the view is inconsistent with Pervasiveness. One might worry my view faces the same difficulty. In cases where one is ignorant or in error about one’s own extended cognitive activities, one will not take an act of conscious thinking as a part of that activity, and so won’t take it as one’s deed. Since this can happen, the view is inconsistent with Pervasiveness.

However, in the first instance, it is the *intention* to engage in an extended cognitive activity that explains the practical character of one’s awareness of one’s action (Gibbons 2010), not *knowledge* of one’s intention. Inner hearing has a practical character because it is informed by one’s intentions. Perhaps in the typical case one will in fact possess practical knowledge, since lapses in that sort of knowledge may be rare. But the absence of self-knowledge needn’t interfere with the sense of agency, which is explained by the governing intention. Thus, the present proposal doesn’t face the same objection as the agency as endorsement model.

<sup>17</sup> You might think such cases of “idle active drift” are not subject to Anscombe’s special sense of the question “why?” where an answer gives the agent’s reasons (1957). That’s because it isn’t clear that such actions are performed for reasons at all, at least normative practical reasons. Again, O’Shaughnessy is on to something when he writes that: “just why these rather than those intentions arose as the process advanced is a matter upon which one may lack authority and insight” (2000, 219). However, if answering the question why involves citing the activity of which a particular stretch of thinking formed is a part, then it would seem that this question can be answered even in cases of daydreaming. (See Thompson 2007 on such “naïve” action explanation.)

<sup>18</sup> Thanks to an anonymous referee for raising this issue.

## 6 Conclusion

There is good reason to think that conscious thinking is a manifestation of our agency and that we ordinarily ascribe it to our agency. This paper has sought to explain what justifies us in these attributions. Why does it seem that conscious thinking is something we do rather than something that happens to us? If we reject inflationary accounts, then we cannot simply sift through the phenomenology to identify a particular quale of agency. And if we take seriously claims about the sense of agency, including what I called Campbell's Thesis, we will recognize that whatever underlies the sense of agency is a pervasive feature of our conscious cognitive lives. The agency as endorsement model fails because the feature it identifies is not pervasive: that feature is absent in cases where the sense of agency is not. Instead, I have proposed that the sense of agency for a particular act of conscious thinking is underwritten by intentions to engage in conscious activities like answering a question, remembering a vacation, or imagining an interlocutor. Our acts of conscious thinking are recognizably our deeds because they are recognizably a part of things that we know that we are up to.

**Acknowledgements** Thanks to Anita Avramides, Christian Coseru, an audience at St Hilda's College, Oxford, and two anonymous referees for helpful feedback on this paper.

**Funding** This article was supported by the project "International mobilities for research activities of the University of Hradec Králové", CZ.02.2.69/0.0/0.0/16\_027/0008487.

## References

- Anscombe, G. E. M. (1957). *Intention*. Oxford: Blackwell.
- Bayne, T. (2008). The phenomenology of agency. *Philosophy Compass*, 3(1), 182–202.
- Bermudez, J. L. (2011). Bodily awareness and self-consciousness. In S. Gallagher (Ed.), *Oxford handbook of the self* (pp. 157–179). Oxford: Oxford University Press.
- Bermudez, J. L. (2019). Bodily ownership, psychological ownership, and psychopathology. *Review of Philosophy and Psychology*, 10(2), 263–280.
- Bortolotti, L. (2009). *Delusions and other irrational beliefs*. Oxford: Oxford University Press.
- Campbell, J. (1999). Schizophrenia, the space of reasons, and thinking as a motor process. *The Monist*, 82(4), 609–625.
- Davidson, D. (1980). *Essays on actions and events*. Oxford: Blackwell.
- Davies, M., Coltheart, M., Langdon, R., & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry, and Psychology*, 8(2/3), 133–158.
- Doyle, C. (2018). Agency and observation in knowledge of one's own thinking. *European Journal of Philosophy*, 27(1), 148–161.
- Fernandez, J. (2010). Thought insertion and self-knowledge. *Mind Language*, 25(1), 66–88.
- Fernandez, J. (2013). *Transparent minds*. Oxford: Oxford University Press.
- Fernyhough, C. (2016). *The voices within: The history and science of how we talk to ourselves*. London: Basic Books.
- Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1), 5–20.
- Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove: Erlbaum (UK) Taylor & Francis.
- Gauker, C. (2018). Inner speech as the internalization of outer speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices* (pp. 53–77). Oxford: Oxford University Press.

- Gibbons, J. (2010). Seeing what you're doing. *Oxford Studies in Epistemology*, 3, 63–85.
- Graham, G., & Stephens, L. (2000). *When self-consciousness breaks*. Cambridge, MA: MIT Press.
- Hoerl, C. (2001). On thought insertion. *Philosophy, Psychiatry and Psychology*, 8(2/3), 189–200.
- Hulburt, R. T., Heavey, C., & Kelsey, J. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, 22, 1477–1494.
- Hurlburt, R. T., & Akhter, S. A. (2006). The descriptive experience sampling method. *Phenomenology and the Cognitive Sciences*, 5, 271–301.
- Hulburt, R., & Heavy, C. (2018). Inner speech as pristine inner experience. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices*. Oxford: Oxford University Press.
- Irving, Z. (2016). Mind-wandering is unguided attention: Accounting for the 'purposeful' wanderer. *Philosophical Studies*, 173, 547–571.
- Jaspers, K. ([1913] 1997). *General psychopathology*. trans. J. Hoening & M.W. Hamilton. Baltimore, MD: Johns Hopkins University Press.
- Langland-Hassan, P. (2018). From inner speech to essence: The auditory nature of inner speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices* (pp. 78–104). Oxford: Oxford University Press.
- Leite, A. (2018). Changing one's mind: Self-conscious belief and rational endorsement. *Philosophy and Phenomenological Research*, 97(1), 150–171.
- Machery, E. (2018). Know thyself: Beliefs vs desires in inner speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices* (pp. 261–275). Oxford: Oxford University Press.
- Martin, M. G. F. (1995). Bodily awareness: A sense of ownership. In J. L. Bermúdez, A. Marcel, & N. Eilan (Eds.), *The body and the self* (pp. 267–289). Cambridge, MA: MIT Press.
- McDowell, J. (2011). Some remarks on intention in action. *The Amherst Lecture in Philosophy*, 6, 1–18.
- Mellor, C. H. (1970). First rank symptoms of schizophrenia. *British Journal of Psychiatry*, 117(536), 15–23.
- Metzinger, T. (2013). The myth of cognitive agency: Subpersonal thinking as a cyclically recurring loss of mental autonomy. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2013.00931>.
- Moran, R. (2001). *Authority and estrangement*. Princeton, NJ: Princeton University Press.
- Moran, R. 2004. Anscombe on 'practical knowledge'. In *Agency and Action*, Royal Institute of Philosophy Supplement 55: 43–68. Cambridge: Cambridge University Press.
- Musholt, K. (2015). *Thinking about oneself: From nonconceptual content to the concept of a self*. Cambridge, MA: MIT Press.
- Nichols, S., & Stich, S. (2001). *Mind-reading*. Oxford: Clarendon Press.
- O'Shaughnessy, B. (2000). *Consciousness and the world*. Oxford: Oxford University Press.
- O'Brien, L. (2007). *Self-knowing agents*. Oxford: Oxford University Press.
- O'Brien, L. (2013). Obsessive thoughts and inner voices. *Philosophical Issues*, 23, 93–108.
- Peacocke, C. (1999). *Being known*. Oxford: Oxford University Press.
- Pickard, H. (2010). Schizophrenia and the epistemology of self-knowledge. *European Journal of Analytical Philosophy*, 6(1), 55–74.
- Proust, J. (2009). Is there a sense of agency for thought? In L. O'Brien & M. Soteriou (Eds.), *Mental actions* (pp. 253–279). Oxford: Oxford University Press.
- Roessler, J. (2013). Thought insertion, self-awareness, and rationality. In K. Fulford, et al. (Eds.), *Oxford handbook of philosophy and psychiatry* (pp. 658–672). Oxford: Oxford University Press.
- Roessler, J., & Eilan, N. (Eds.). (2003). *Agency and self-awareness: Issues in philosophy and psychology* (pp. 383–405). Oxford: Clarendon Press.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Ryle, G. (2009). *Collected essays 1929–1968*. London: Routledge.
- Saks, E. (2007). *The centre cannot hold: A memoir of my schizophrenia*. London: Virago.
- Scheer, J. (2009). Experience and self-consciousness. *Philosophical Studies*, 144, 95–105.
- Schwenkler, J. (2011). Perception and practical knowledge. *Philosophical Explorations*, 14(2), 137–152.
- Setiya, K. (2007). *Reasons without rationalism*. Princeton University Press.
- Shepherd, J. (2019). Why does the mind wander? *Neuroscience of Consciousness*, 1, 1–9. <https://doi.org/10.1093/nc/niz014>.
- Soteriou, M. (2013). *The mind's construction*. Oxford: Oxford University Press.
- Thompson, M. (2007). *Life and action*. Cambridge, MA: Harvard University Press.
- Thornton, T. (2007). *Essential philosophy of psychiatry*. Oxford: Oxford University Press.

- Velleman, J. D. (1989). *Practical reflection*. Princeton, NJ: Princeton University Press.
- Wilkinson, S., & Ferynhough, C. (2018). When inner speech misleads. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices* (pp. 244–260). Oxford: Oxford University Press.
- Zahavi, Dan. (2005). *Subjectivity and selfhood: Investigating the first-person perspective*. Cambridge, MA: MIT Press.
- Zahavi, D., & Kriegel, U. (2015). For-me-ness: What it is and what it is not. In D. Dahlstrom, A. Elpidorou, & W. Hopp (Eds.), *Philosophy of mind and phenomenology* (pp. 36–53). Abingdon: Routledge.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.