

**Thought Insertion:  
Abnormal Sense of Thought Agency or Thought Endorsement?**

Paulo Sousa & Lauren Swiney\*

**Abstract.** The standard approach to the core phenomenology of thought insertion characterizes it in terms of a normal sense of thought ownership coupled with an abnormal sense of thought agency. Recently, Fernández (2010) has argued that there are crucial problems with this approach and has proposed instead that what goes wrong fundamentally in such a phenomenology is a sense of thought commitment, characterized in terms of thought endorsement. In this paper, we argue that, even though Fernández raises new issues that enrich the topic, his proposal cannot rival the version of the standard approach we shall defend.

**Keywords** Thought Insertion, Sense of Thought Ownership, Sense of Thought Agency, Sense of Thought Endorsement, Self-Knowledge, Folk Psychology, Schizophrenia

## 1. Introduction

Consider the following reports of what is happening with one's mental life:

---

\*Authorship is equal. Address correspondence to Paulo Sousa or Lauren Swiney, Institute of Cognition and Culture, Queen's University, Belfast, 2-4 Fitzwilliam Street, BT71NN, UK, email: [p.sousa@qub.ac.uk](mailto:p.sousa@qub.ac.uk), [lswiney01@qub.ac.uk](mailto:lswiney01@qub.ac.uk).

(R1) Thoughts are put into my mind, like ‘Kill God’. It is just like my mind working, but it isn’t. They are not my thoughts. They belong to this guy, Chris. They are his thoughts. (Frith 1992, p. 66)

(R2) 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)

(R3) I have never read nor heard them; they come unasked; I do not dare to think I am the source but I am happy to know of them without thinking them. They come at any moment like a gift and I do not dare to impart them as if they were my own (Jaspers 1963, p. 123).

These reports indicate a diagnostic symptom of schizophrenia, a delusion called thought insertion, the topic of this paper.<sup>1</sup>

The classification and investigation of delusions has a long cross-disciplinary history (Jaspers 1963; Oltmanns and Maher 1988; Garety and Hemsley 1994, Radden 2010). Thought insertion, in particular, is one of the most fascinating of the so-called “bizarre delusions” (DSM-IV-TR). Given its status as a *first-rank* symptom of schizophrenia, which implies that the regular occurrence of the symptom over a month can alone justify a diagnosis

---

<sup>1</sup> We shall assume that delusions are belief-like attitudes (see DSM-IV; Bayne & Pacherie 2005; for a different position, see Currie 2000).

of schizophrenia, the study of thought insertion has significant practical implications. Moreover, patients' insistence that they are experiencing a thought that is in some sense "not theirs" seems to violate basic intuitions about the relation between self, mind and thought. Thus, the study of thought insertion also sheds light on our basic phenomenological presuppositions.

Discussions of thought insertion may deal with three distinct levels of analysis: the characterization of the common behavioral features (e.g. features of verbal reports) indicating that patients have had a similarly odd experience with their thoughts; the characterization of the basic features of the experience of thought insertion, which may include both feelings that constitute the qualia of the experience and judgments that constitute the conceptualization of the experience;<sup>2</sup> the characterization of the causal factors leading patients to undergo such an experience. The first level provides evidence supporting the existence of a phenomenology of thought insertion; the second a description of the core structure of such a phenomenology; the third a theoretical explanation of its occurrence. Although we don't claim that these levels are completely independent of each other (e.g. the characterization of the evidence or of the phenomenology may be somewhat theory-laden), we accept that they have some autonomy.

Over the last number of years, researchers from various theoretical convictions have claimed that the combination of a normal sense of thought ownership with an abnormal sense of thought agency constitutes the core phenomenology of thought insertion (see references below). Although there is some variation in the understanding of this phenomenological core, some of which will be explicated in what follows, this can be considered the standard approach to the subject. However, Fernández (2010) has recently argued that there are critical shortcomings with the standard approach and has put forward an alternative proposal where the core phenomenology of thought insertion is characterized in terms of an abnormal sense

---

<sup>2</sup> We use the terms "sense," "experience" and "phenomenology" with the broader meaning that subsumes both feelings and judgments (Synofzik, Vosgerau & Newen 2008).

of thought commitment. Furthermore, Fernández's characterization leads to an aetiological account that is quite distinct from those related to the standard approach—if he is right, a range of explanations connected to the standard approach should be discarded.

The aim of this paper is to vindicate a version of the standard approach, arguing that Fernández's proposal falls short of furnishing a better alternative. For the most part we leave the explanatory level of the discussion aside, focusing instead on the aforementioned first two levels—the proper characterization of the core phenomenology and its support by the available evidence. The focus on the characterization of the explanandum and the different aspects of the supporting evidence is salutary because current approaches are often exceedingly theoretically driven and don't clearly show how their characterization of the phenomenology relates to the details of the evidence. Furthermore, our overall discussion will set empirical constraints that any theoretical account should meet in order to constitute a plausible explanation.

In the next section, we concentrate on our version of the standard approach and the different aspects of the available evidence. In the following one, we assess Fernández's positive contribution. We conclude by briefly addressing Fernández's principal criticisms to the standard approach.

## **2. Our version of the standard approach**

In this section, we first outline what we take to be the best characterization of the basic phenomenology of thought insertion. Then, we describe the different aspects of the evidence, showing in detail how they support our characterization. While our characterization

draws on the standard approach to the topic, it incorporates novel elements that significantly improve this approach.

## **2.1. Sense of thought ownership without a sense of thought agency**

A well-known way of characterizing the basic phenomenology of thought insertion is to say that, in thought insertion, a normal sense of thought ownership is combined with an abnormal sense of thought agency (Campbell 2002; Carruthers 2007; Gallagher 2004; Gerrans 2001; Proust 2006; Graham and Stephens 2000; Young 2007).<sup>3</sup>

There are two slightly different notions of the sense of thought ownership, although this point hasn't been properly recognized by the current literature (for exceptions, see Billon, *forthcoming*; Fernández 2010; also Gallagher 2000, note 4). One notion, advocated paradigmatically by Stephens and Graham, is discussed via the inseparability thesis:

In being aware of my thoughts I am, ipso facto, aware of them as occurring within the boundary of me. Hereafter we shall term the claim that the sense of subjectivity is a proper part of introspection the "inseparability thesis," since the idea is that introspection and subjectivity form a single, integrated whole. (Stephens and Graham 1994, p. 2)

Another, advocated paradigmatically by Campbell, is framed by the issue of immunity to error through misidentification:

---

<sup>3</sup> Some of these authors use the term "ownership" with a superordinate meaning that subsumes both senses to be characterized here (e.g. Campbell, Carruthers), and some use the term "authorship" to refer to the sense of agency (e.g. Gerrans, Young).

Where does this leave the notion of immunity to error through misidentification of present-tense self-ascriptions of thoughts? (...) there is a sense in which the possibility of introspective knowledge by you is enough to make it a thought of yours. There is just a contradiction in supposing that a thought I am introspecting is not being introspected by me. (Campbell 2002, p. 36-7)

The immunity question is whether an introspected thought is always experienced *as experienced by me*, while the inseparability question is whether an introspected thought that is experienced as experienced by me is always experienced *as located in me*. Thus, Stephens & Graham's sense presupposes Campbell's, but not the reverse, a dissociation that perhaps occurs in certain cases of multiple-personality disorder, where one self might experience a thought as located in another self.

In the context of characterizing the basic phenomenology of thought insertion, Stephens & Graham's notion is the most relevant because its additional component of subjectivity boundary is fundamental to capture some aspects of such a phenomenology. Accordingly, we characterise the sense of thought ownership as the sense that the thoughts that I introspect occur within the boundaries of my subjectivity. For instance, if I start to imagine the place where I'm going to spend my vacation and become aware of the resultant images, I have a sense that these images are occurring in my inner space, namely, that they are part of my stream of consciousness.<sup>4</sup> There are two distinct aspects of our characterization here. The first is the process of introspection by which I experience the thoughts.<sup>5</sup> The second is the

---

<sup>4</sup> Some authors have criticized this sense of ownership on the grounds that it uses a misleading spatial metaphor (e.g. Bortolotti and Broome 2009; Thornton 2002). However, metaphors may sound misleading if they are understood too literally. For us, the important question is whether the notion of *occurrence within limits* can be given a plausible psychological interpretation, when the theoretical aim is one of a phenomenological description concerning the experience with thoughts. Since plausibility in this respect has to do with being faithful to how people experience thoughts, and since people do seem to regard occurrent thoughts as happening within internal mind-personal limits ("inner space", "stream of consciousness"), we do not see any fundamental problem with the spatial metaphor, properly interpreted.

<sup>5</sup> In this article, we shall assume that introspection exists for at least some types of thoughts (see Carruthers

localization of the experienced thoughts within the boundaries of my subjectivity. In our view, there is a normal sense of thought ownership in thought insertion: the patients experience the thoughts as occurring in their subjectivity.

There are two broadly distinct notions of thought agency. Take two cases in which the images of the place where I'm going to spend my vacation may occur in my subjectivity: (i) I'm concentrated on imagining the place where I'm going to spend my vacation and, consequently, images of such a place appear in my stream of consciousness; (ii) while concentrated on other matters, unsolicited images of the place I'm going to spend my vacation appear in my stream of consciousness. In the first case, I have a sense that I have formed the intention (or, at least that I'm trying, with some effort or not) to imagine and I have a general sense of control over the direction of my thoughts. In the second case, I have neither a sense of having formed an intention nor a sense of trying to imagine (hence, no sense of effort either). Moreover, I have no general sense of directional control.

The first notion of agency, related to the first case of imagining, is that of agent intentional guidance—I have a sense that I'm intentionally guiding the occurrence of the images. The second notion of agency, related to both cases of imagining, is that of agent causation—I have a sense that I'm the causal source producing the occurrence of the images, whether intentionally or not. In other words, the first notion of agency corresponds to the subordinate concept of intentional action, which relates only to the intentional production of thoughts; the second notion corresponds to the superordinate concept of action, which subsumes both the intentional and the unintentional production of thoughts, and is to be analysed in terms of the agent causation of thoughts.

While current versions of the standard approach draw from some notion of intentional guidance when characterizing the sense of thought agency relevant to the phenomenology of

---

2009a for a discussion on the extent to which there in fact exists introspective access to our mental life or this is just a false assumption built in our folk psychology).

thought insertion, we draw from the superordinate concept of action.<sup>6</sup> Accordingly, we characterise the sense of thought agency as the sense that I'm producing the thoughts occurring in my subjectivity, i.e., that I'm bringing about the occurrence of these thoughts, intentionally or not. In our view, there is an abnormal sense of thought agency in thought insertion. There are two distinct aspects of this abnormal sense. The patients do not experience the occurrence of the thoughts in their subjectivity as being produced by them. Moreover, the patients experience this occurrence as being produced by an alien agent.

In sum, the normally fused senses of thought ownership and thought agency, as characterized, are dissociated in thought insertion due to an abnormal sense of thought agency.

## 2.2. The different aspects of the evidence

The most important evidence comes from reports like those cited in the introduction. Four separate parts of these reports support our characterization of the basic phenomenology of thought insertion.

The first, which we shall call the *locus-of-occurrence* part, describes the “inserted” thoughts in terms of where they occur, indicating that the patients had a normal sense of thought ownership: (R1) *Thoughts are put into my mind ...;* (R2) ... the *thoughts* of Eamonn

---

<sup>6</sup> It is important to emphasize that our discussion here concerns the sense of agency for thought, not whether there are in fact mental actions or intentional actions. For discussions focusing more, though not exclusively, on the extent to which actions or intentional actions are part of our mental life, see Buckareff 2005; Carruthers 2009b; Mele 2009; Peacocke 2007; Strawson 2003. For a more complete analysis of the superordinate folk concept of action involved in our sense of thought agency, see Jackendoff, 1990, ch. 7. For a more complete discussion of components that may be involved in the sense of thought intentional guidance, see Gallagher in press; Pacherie 2008.

Andrews *come into my mind*. The remaining parts relate to the abnormal sense of thought agency.

The second, which we shall call the *absence-of-self-agency* part, shows that the patients do not have a sense of being the producers of the thoughts they become aware of: (R1) *It is just like my mind working, but it isn't.*; (R3) ... *I do not dare to think I am the source* but I am happy to know of them *without thinking them*.<sup>7</sup> The third, which we shall call the *alien-agency* part, describes some process of insertion of thoughts in the patients' mind, indicating that the occurrence of the thoughts in their stream of consciousness wasn't experienced as being produced by them but by another agent: (R1) *Thoughts are put into my mind ...*; (R2) *He treats my mind like a screen and flashes thoughts onto it like you flash a picture.*

The third part evinces that the sense of alien thought agency supposes a more complex action of production than the one related to the normal sense of thought agency—the alien agent produces the thoughts in his subjectivity *and* makes them occur in the patient's subjectivity. Moreover, this “and” may be understood in terms of simultaneity or sequence. In terms of simultaneity, while producing the thoughts in his subjectivity, the alien agent makes them occur in the patient's subjectivity, as if the alien agent were thinking in the patient's subjectivity. In terms of sequence, the alien agent produces the thoughts first and then makes them occur in the patients' subjectivity, a succession that may be mediated by all kinds of devices supposed to transmit thoughts. Finally, the alien agent in itself may be an entity other than a person, being anthropomorphized at least to the point of being capable of producing thoughts, as indicated by this report:

---

<sup>7</sup> The expression “my mind working” in this passage of R1 uses “mind” differently from the usage of “mind” in the expression “into my mind” described earlier in relation to the sense of thought ownership—my mind-working is a system whose activities produce my thoughts instead of a phenomenological space where my thoughts occur. We take this notion of mind production to be roughly equivalent to the notion of agent production.

(R4) ... the things that come are not the things that I have been thinking about ...

They [the household appliances] kind of short circuit the brain, and bring their message (Jackson and Fulford 1997, p. 45).<sup>8</sup>

The fourth part, which we shall call the *thought-myness* part, is related to a specific interpretation of the meaning of the language of thought ownership utilized in the reports—i.e. the meaning of words like “my” or “belong” as applied to *thoughts*: (R1) They are *not my thoughts*. They *belong to this guy* ...; (R2) ... the *thoughts of Eamonn Andrews* come into my mind. There are *no other thoughts* there, *only his* ...; (R3) ... I do not dare to impart them *as if they were my own*.

To fully understand this point, it is important to keep in mind that the language of ownership in general is semantically underspecified and can convey an indefinite variety of meanings depending on the context. Take the sentence “this is not *my book*.” Depending on the context in which it is enunciated, it may convey that this is not *the book that I possess*, or *the book that I wrote*, or *the book about me*, or *the book that I prefer*, or *the book that I'm thinking of* etc. To use the vocabulary of contemporary lexical pragmatics (Wilson and Carston 2007), in each of these contexts an ad hoc concept of myness is intended and conveyed, each portraying a quite different relation between the book and the I.

Now, when the patients claim that the “inserted” thought is not theirs (or belongs to an alien agent), the interpretation of our standard approach cannot be that the language of thought-myness is meant to convey the idea of ownership related to the sense of thought

---

<sup>8</sup> This report is given by a subject without a diagnosis of schizophrenia, but whose religious experiences involved delusions and hallucinations as defined by the Present State Examination (Wing 1974)—see Jackson and Fulford, 1997, p. 43.

ownership characterized earlier: they are not saying that they had experienced a thought as occurring outside their subjectivity (or inside the subjectivity of another agent). After all, the patients state that the thoughts occur in their minds.<sup>9</sup>

According to our interpretation, the thought-myness part is not meant to convey an idea of right-of-possession similar to the first interpretation of the book example above either: the patients are not saying that they had experienced a thought that they feel they do not possess (or feel another agent possesses). Although some authors seem to imply that the idea of right of possession is relevant to the phenomenology of thought insertion (e.g. Langland-Hassan 2008), and others try to make an explicit case for it (e.g. Bortolotti and Broome 2009), this is misleading. Take the “inserted” thought “Kill God” explicated by R1. With the thought-myness part of the report, the patient is not trying to convey any absent sense of right of possession (or the alien sense of right of possession) over this thought.

In fact, what our standard approach claims is that the thought-myness part is, in most cases, meant to convey the idea of a relation between the inserted thought and the I concerning the sense of agency: when the patients say that the thought is not theirs, they are saying that they are not the producers of the thought they experienced; when they say that the thought belongs to another agent, they are saying that another agent produced the thought they experienced.

It should be apparent that we assume a fully deflationary view of the thought-myness part of the evidence. Some authors have criticized proponents of the standard approach for denying the face-value, literal meaning that is meant to be conveyed by the thought-myness part of reports, and for reinterpreting it in terms of a sense of agency (*cf.* Billon *forthcoming*;

---

<sup>9</sup> Note that the language of ownership when applied to the mind (“Thoughts come in *my mind*”) instead of to specific thoughts, as in the thought-myness part (“*Thoughts are not mine*”), seems to mean simply a part-whole relation between the mind and the I (Thoughts come in *the mind that is a constitutive part of me*).

Bortolotti and Broome 2009). However, the assumption of a face-value, literal meaning meant to be conveyed supposes a quite naive understanding of the semantics and pragmatics of the language of ownership and the charge of reinterpretation is therefore misleading. Since the language of ownership can convey an indefinite variety of meanings depending on the context, the thought-myness language in the context of thought-insertion reports requires a specific interpretation (not the reinterpretation of a literal meaning), one whose support depends completely on the other aspects of the evidence.<sup>10</sup> In reality, these other aspects indicate that the thought-myness part of the evidence may often only serve to emphasize the information already conveyed by the absence-of-self-agency and alien-agency parts of reports. For instance, take this passage: (R1) Thoughts are put into my mind, like “Kill God”. It is just like my mind working, but it isn’t. They are *not my thoughts*. The patient says that it is not his mind that is producing the thoughts and is using the language of thought ownership just to rephrase and emphasize this point. Or take this passage: (R3) I do not dare to think I am the source but I am happy to know of them without thinking them. They come at any moment like a gift and *I do not dare to impart them as if they were my own*. Again, the patient is simply emphasizing via the language of thought ownership that she does not have the sense of being the producer (“source”) of the thoughts.

The clinical literature has recently put emphasis on the thought-myness part of reports for the diagnosis of thought insertion (the patients say “thoughts are not mine or belong to another agent”) and even for characterising its core phenomenology (in thought insertion the patients do not feel that they own the thoughts) (see Koehler 1979; Mullins and Spence 2003; O’Grady 1990). However, when one inspects this literature’s interpretation of the language of

---

<sup>10</sup> Authors that suppose such a face-value meaning do not provide any clear independent evidence for it. Often, their specific characterization of the content of the face value meaning is informed simply by their own theoretical assumptions. For example, Billon (*forthcoming*) supposes that the thought-myness part of reports is meant to convey literally that the patients do not have a sense of subjectivity (characterized in terms of phenomenal consciousness) for the “inserted” thought, but does not provide any clear independent evidence for this interpretation other than his theoretical assumption that the core phenomenology of thought insertion is to be characterized in terms of an abnormal sense of subjectivity.

thought ownership, it aligns with our interpretation:

Thinking, like all conscious activity, *is experienced as an activity which is being carried out by the subject...* There is *a quality of “my-ness” connected with thought*. In schizophrenia this sense of the possession of one’s thoughts may be impaired and the patient may suffer from alienation of thought... [The patient] is certain that alien thoughts have been inserted in his mind. (Fish 1976, p. 48; emphases are ours)

Alienated experiences were described as the patient’s awareness that thoughts (...) were NOT HIS OWN *in the sense that they were coming from an outside source*. (Koehler 1979, p. 241; capitals are in the original; emphasis is ours)

Furthermore, despite the emphasis, the clinical literature describes cases of thought-insertion reports in which the language of ownership is clearly *not* present, implying that the thought-myness part is not fundamental. Take the following:

(R5) One evening the thought was given to me electrically that I should murder Lissi.  
(Mullins and Spence 2003 p. 295).

Also, take this passage from R4: They [the household appliances] kind of short circuit the brain, and bring their message. Does the expression “their message” constitute a thought-myness part of the report? If so, it seems to mean simply *the message they want to convey*, which does not add much in terms of evidence. If not, then we have another example of a thought-insertion report in which the thought-myness part is not present.

Finally, while thought-myness has been emphasized in the clinical literature to distinguish thought insertion from the similar symptom of thought *control*, both in terms of the evidence (patients with thought control do not say that “thoughts are *not mine* or *belong to another agent*”) and in phenomenological terms (patients with thought control still feel that they *own* the thoughts), we claim that our account provides a clearer distinction between the two, without relying on thought-myness. Thus, in terms of the evidence, what first specifies thought-insertion reports in relation to thought-control reports is the presence of the absence-of-self-agency or alien-agency parts of reports, as we characterized them—the thought-myness part is not necessary and it cannot alone provide such a specification. In terms of core phenomenology, we suggest that while in thought insertion what is abnormal is the sense of agent causation, in thought control what is compromised is the sense of agent intentional guidance—there is no need to use the language of thought-myness in the characterization since this language would have to be interpreted in terms of the sense of agent causation anyway.

Before we move to Fernández’s alternative proposal, it is important to point out something in our discussion that we have left deliberately underspecified. We have been using the words “sense”, “phenomenology” and “experience” with their broader meaning that includes both feelings and conceptual judgments (see note 2), without discussing the extent to which or in which way feelings are related to conceptual judgments and vice-versa—e.g. is there a feeling of absence of self-agency attached to the judgment of absence of self-agency? If so, is this feeling the cause of this judgment or is it rather one of its consequences? Our characterization is compatible with a whole spectrum of positions—from two-factor accounts defending that feelings are behind some of the judgments (e.g. the one of absence of self-agency) while other judgments (e.g. that a household appliance is prompting the insertion) are just explanatory improvisation, to complete top-down accounts saying that feelings do not

enter into the picture at all or enter into the picture at most as an epiphenomenal effect of judgments. To delineate our position in this complex issue would lead us into aetiological discussions that would move us from the main aim of the paper, which is to discuss the answer to the what-question of the phenomenology of thought insertion given the available evidence.

### **3. Fernández’s proposal**

In this section, first we present Fernández’s proposal in the context of his overall theoretical framework. Then, in turn, we evaluate it regarding the two varieties of evidence he invokes—thought-insertion reports and broader aspects of behaviors related to schizophrenic delusions. We shall show that his proposal is not supported by the evidence and thus cannot rival our standard one.

#### **3.1. Belief awareness without a sense of thought commitment**

In order to understand Fernandez’s proposal, it will be necessary to draw a clear distinction between different types of thought, and to address the question of whether “inserted” thoughts are restricted to thoughts of a specific type. Among other dimensions, thoughts may differ in terms of content (whether they are about vacation or work), coding format (whether they are in a propositional language of thought or in imagery), or “attitude” (if in a propositional language of thought, whether they are beliefs or desires; if in imagery like inner speech, whether they are assertions or commands). For example, one may say to

oneself in inner speech “finish this article!” This thought has a certain content (about finishing an article), coding format (verbal imagery), and attitude (it is a command).

For Fernández, “inserted” thoughts are restricted to beliefs—more specifically, to beliefs of the patients suffering from thought insertion. Thus, Fernández’s proposal is that in thought insertion the patients do not have a sense of thought commitment to some of their own beliefs in that “[they] do not feel pressured to endorse the contents of some beliefs which they are aware of having.” (2010, p. 67)

To see what exactly Fernández has in mind, we need to explicate what, according to him, is the normal way of having a sense of endorsing the content of one’s own beliefs. Fernández defends a bypass version of the transparency model of self-knowledge of beliefs, according to which, when one adopts a first-person perspective upon one’s own beliefs, one knows of a first-order belief  $P$  that one has and forms the meta-belief  $I$  believe that  $P$ , by directing one’s attention outwards, towards the features of the world that provide evidence for  $P$ , instead of by introspection (Fernández 2003).<sup>11</sup> It is for this reason that, in thus forming the meta-belief  $I$  believe that  $P$ , one experiences the feeling that one endorses the content of the first-order belief  $P$ —the epistemic justification for the first-order belief and for the formation of the related meta-belief is the same. For example, suppose that one has a first-order belief with the propositional content *the grass is green* in one’s belief box, that this belief becomes occurrent in one’s working memory, and that one becomes aware of it (but not *qua* one’s own belief). In order to know whether this first-order belief is one’s own belief, one has to bypass the introspective focus by shifting one’s attention to the evidence for it (e.g., by looking at the green grass in one’s scene or by retrieving memories of the grass looking green). It is by taking into account this type of transparent evidence that one forms the meta-belief with the propositional content  $I$  believe that *the grass is green* and

---

<sup>11</sup> For a general discussion of different models of self-knowledge, including Fernández’s, see Gertler 2008.

experiences the feeling that one endorses the first-order propositional content *the grass is green*.

It is important to note that, when talking about bypassing an introspective focus, “introspection” is not being used to mean simply inspecting one’s mind, which may include any kind of memory search and retrieval, but to mean something more specific that involves some sort of meta-representational awareness—i.e. the fact that one searches and retrieves some information from memory does not necessarily imply that one has introspected it. That is why, in denying that one can acquire knowledge of one’s beliefs via introspection, Fernández is not implying that an outward look cannot involve evidential information retrieved from memory. Moreover, Fernández does not (or should not) claim that introspection does not give any kind of access to one’s beliefs—if one simply focuses one’s attention on a first-order belief  $P$  without shifting it to the evidence for  $P$ , one meta-represents  $P$ , but without forming the belief *I believe that P*. That is why, according to his model, one may be aware of one’s own beliefs without knowing them.<sup>12</sup>

Now, for Fernández, it is impairment of the bypass procedure of knowing one’s first-order beliefs that leads to the phenomenology of thought insertion. When the patients become aware of some of their first-order beliefs, they do not form the related meta-beliefs and therefore do not have the feeling of endorsing their first-order content. According to him, this breakage happens because the patients have a hyper-reflexivity tendency that entails an inability to focus on the relevant evidence—instead of attending to transparent evidence for their first-order beliefs, they tend to keep focusing on their own introspective experience. Take the “inserted” thought evinced by this report: (R5) ...the thought was given to me

---

<sup>12</sup> Compare this position with authors, like Nichols and Stich (2003), who understand introspection of beliefs as involving a monitoring mechanism that takes as input a belief  $P$  and delivers as output the belief *I believe that P*, thereby necessarily providing self-knowledge of beliefs. In Fernández’s model, introspection cannot have as output a representation with the content *I believe that P* but only a meta-representation of  $P$  in the abstract (that is, not embedded in *I believe that...*).

electrically that I should murder Lissi. According to Fernández's explanation, the patient has a first-order belief with the propositional content *I should murder Lissi*, but, due to the inability to focus on the evidence for it, the patient does not form the meta-belief *I believe that I should murder Lissi*, and does not get the feeling of endorsing the first-order content.

Before we move to the evidence for Fernández's proposal, it is important to emphasize that the pertinent thought at stake in this discussion (i.e. the thought whose content, according to Fernandez, constitutes a first-order belief of the patient) is the “inserted” thought (e.g. the thought *I should murder Lissi*), not the delusional thought that a thought has been inserted in one's mind (e.g. the thought *I should murder Lissi was given to me electronically*).

### **3.2. Thought insertion reports**

According to Fernández, the structural characteristic of *all* thought insertion reports is a statement of awareness without ownership (“AWO”) in that “for any subject with thought insertion S, there is some belief B such that: (i) S claims to have B; (ii) S claims that B is not his/her belief” (2010, p. 68). These two clauses are Fernández's rendering of the locus-of-occurrence and thought-myness parts of reports we isolated in the section “The different aspects of the evidence” (“Thoughts come into my mind”, “Thoughts are not mine”, respectively).<sup>13</sup> They are supposed to comprise the main evidence supporting his characterization of the core phenomenology of thought insertion. The first would be evidence

---

<sup>13</sup> Remember that the there are two sides of the thought-myness part of reports (thoughts are not mine and thoughts belong to another agent). Fernández's second clause incorporates only one of these sides because his approach is intended to capture first and foremost what he takes to be the feeling component of the phenomenology of thought insertion (i.e. the patient's feeling of non-endorsement of the content of the “inserted” thought) and it is the claim that a thought is not mine that is related to this feeling. This qualification does not affect in any respect our following criticisms.

that the patients become aware of their beliefs.<sup>14</sup> The second would be evidence that the patients do not have a feeling of endorsing the content of these beliefs: “[The patient] does not experience those beliefs as forcing any particular picture of the world on him. This is why the patient affirms that those beliefs are not his. After all, he does not feel compelled to endorse their contents.” (2010, p. 78)

The first critical point we would like to raise relates to Fernández’s structural characterization of reports in terms of the patients making claims *concerning their first-order beliefs*. For this characterization to constitute a reasonable depiction of the evidence, he would have to defend that the parts of reports where the patients make explicit the content of the “inserted” thoughts indicate that these thoughts are always beliefs (never other types of thoughts) and always beliefs of the patients. However, the parts of reports that explicate the content of “inserted” thoughts do not support the idea that patient’s claims about “inserted” thoughts are confined to claims about (their) beliefs.

Take the “inserted” thought evinced by this report: (R1) Thoughts are put in my mind, like ‘*Kill God*’. Fernández supposes that the original thought being reported was a first-order belief of the patient with the propositional content *I ought to kill God*. But, given that the patient reported the “inserted” thought as a command, a more plausible interpretation is that the original thought was a command in inner speech imagery.<sup>15</sup> Anyway, if *I ought to kill God* were a first-order belief of the patient, what could be the type of evidence that the patient

---

<sup>14</sup> It worth noticing that it is misleading for Fernández to phrase his first clause in terms of the patient claiming to have a belief, or, as in other parts of the article, in terms of the patient making a claim like “I believe that such-and-such” (2010, p. 71). What Fernández has in mind, as is suggested by his naming his first clause in terms of *awareness*, is simply that the patients claim to have become aware of their beliefs in the way that we characterized his model above—i.e. aware of their beliefs but not *qua* their beliefs.

<sup>15</sup> There are a variety of positions in the literature regarding the types of thoughts that might be considered to be inserted, though none shares Fernandez’s position. Some exclude thoughts in imagery format (e.g. inner speech) but beyond that allow that all propositional thoughts, beliefs or not, could be “inserted” thoughts (Vosgerau and Newen 2007). Others restrict cases of thought insertion to alienated inner speech (Langland-Hassan 2008). Still others talk primarily but not exclusively about inner speech (Graham and Stephens 2000). This issue is also important in deciding the degree to which thought insertion and auditory verbal hallucinations (AVHs) share a common phenomenology. We think that the current evidence does not indicate that “inserted” thoughts are restricted to specific types of thoughts and we are sympathetic to the idea that one should not make a radical distinction between thought insertion and AVHs.

is not attending to and would constitute transparent ground for *I ought to kill God*, and, if *ought* implies *can*, also for *I can kill God*? True, the patient may not be shifting his attention to the fact that, even though she has been praying regularly and hasn't done anything wrong, many bad things are happening in her life. This could be the basis of a belief that, in not protecting her, God deserves punishment, and even of a desire that God should die. However, this is still quite far away from what could constitute the basis of the belief *I ought to kill God*.

Or take Fernández's interpretation of this passage of this report: (R2) 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. Fernández claims that the thoughts *the garden looks nice* and *the grass looks cool* are being reported as "inserted" thoughts and that these are first-order perceptual beliefs the patient became aware of at that point, but not as beliefs of hers. However, this passage does not show that the patient is reporting these thoughts as "inserted" thoughts. For us, the best interpretation, indicated by the conjunction "but", is that the patient is reporting the thoughts she formed when she looked at the garden, which were interrupted by the "inserted" thoughts, whose content the patient does not specify at all in the report. Our interpretation is also consistent with the fact that "inserted" thoughts are typically considered to be surprising in the context of what the patient is currently thinking (see R4, in particular).

Thus, Fernández's characterization of the evidence is an unjustified, overly theory-laden depiction of the evidence insofar as it interprets patient's claims about thoughts more generally in terms of claims about beliefs (of the patients). Having to count heavily on such a delimitation of "inserted" thoughts, without providing any solid support for it, Fernández's proposal already loses much of its motivation.

The second critical point we would like to raise concerns Fernández's second evidential clause. To begin with, it is simply not true that a claim that the "inserted" thought

is not mine is a structural feature of thought insertion reports, since there are thought-insertion reports that do not contain such a claim (see R5, which, ironically, Fernández cites in his article). More importantly, even if Fernández were to make an ad hoc move and exclude, by definition, such reports from the realm of thought-insertion reports, the interpretation of the thought-myness part of reports entailed by his proposal cannot stand, as we shall show.

Fernández's proposal entails that what the patients are trying to convey via the thought-myness part is a sense of non-endorsement of the content *P* of the “inserted” thought. More specifically, this sense of non-endorsement has to be a sense of neutrality instead of a sense of disagreement concerning *P*—by saying that a thought with a content *P* is not hers, the patient can be saying only that *She does not believe that P*, not that *She believes that not P*. This is because, given Fernández's theoretical framework, forming the belief *I believe that not P* and having the concomitant feeling of disendorsing *P* from a first-person perspective would have to come from the same bypass procedure that checks evidence for *P*, but this procedure is impaired.

However, many reports suggest that the patients are not neutral about the “inserted” thoughts. On the one hand, some reports indicate “inserted” thoughts that are positive or quite banal (“good job”, “ok”), suggesting that the patients probably endorse many of them (see Graham and Stephens 2000, p. 169; Taylor and Heiser 1971; Vosgerau and Newen 2007). On the other hand, some reports indicate “inserted” thoughts that are threatening to the patients’ self-image (e.g. R1: ... ‘Kill God’ or R5: ... I should murder Lissi), suggesting that the patients probably disagree with many of them. Either instance undermines Fernández’s interpretation of the thought-myness part of reports, and, ultimately, his characterization of the

phenomenology of thought insertion, which entails this interpretation.<sup>16</sup>

Even if Fernández were to deny such counter-evidence by saying that these supposed cases of endorsement and disagreement are dubious, he would have to provide some additional evidence for his interpretation of the thought-myness part of reports—as we’ve argued, the support for the interpretation of this part depends completely on the other aspects of the evidence. Without such additional evidence, Fernández’s proposal becomes viciously circular, since his interpretation would have to count only on his own theoretical assumptions.

Finally, we would like to discuss two additional reports, some details of which (i.e. parts other than his two clauses), Fernández claims, would give such an additional support to his proposal (thus perhaps for his interpretation of his second clause):

(R6) [H]e said, ‘... it’s like a thought as it comes in ... a thought is very light really, inspirational ... it’s a light feeling where you feel as though I’m actually thinking it ... or you’re receiving it rather ... it’s just a thought but it feels logical say ... it feels pretty normal or fits with what I suspect, [I] wonder if that’s me ... it felt like a piece of information.’ Later he went on: ‘... you find it strange when some different little thought filters through ... why did I think that at this time of day?’ He said you judge it and say, ‘I don’t think that was mine ... you can differentiate.’ (Hoerl 2001, p. 189-90)

---

<sup>16</sup> We would like to acknowledge that it is plausible to suppose that the thought-myness part of reports may be used to convey a non-endorsement meaning when the content of the “inserted” thought is not endorsed—in particular, when the patient disagrees with it. But this is not incompatible with our position, for linguistic expressions may be used to convey more than one meaning. By saying that the “inserted” thoughts are not theirs, the patients may intend to convey both that they are not the producers of and that they do not endorse the ‘inserted’ thoughts, but the former meaning constitutes what they want to say explicitly while the latter would constitute what they want to implicate (see Carston 2002; cf. Grice 1991).

(R7) [S]he said that sometimes it seemed to be her own thought ‘... but I don’t get the feeling that it is.’ She said her ‘own thought might say the same thing ... But the feeling it isn’t the same ... the feeling is that it is somebody else’s ... ’ She was asked if she had other people’s thoughts put inside her head. She said ‘... possibly they are but I don’t think of them in that way ... they were being put into me into my mind ... very similar to what I would be like normally.’ (Hoerl 2001, p. 190)

Fernández claims that the fact that R6 says that the “inserted” thought feels simply like “a piece of information” is quite consistent with his being aware of his beliefs with a feeling of neutrality about their content: “You would expect them [the patients] to speak of it [the “inserted” thought] as being similar to one’s awareness of a representation or a piece of information, which is neutral on whether the introspected state matches the world or not.”(Fernández 2010, p. 79) And he claims that the fact that R6 and R7 say, respectively, that the “inserted” thought “fits with what I suspect” and that her “own thought might say the same thing ...” is a way of expressing the contingent coincidence between what they introspect and what they believe, given the failure of the bypass procedure: “Nothing in their awareness of having a certain belief seems to force them to endorse the content of that belief. It seems natural, then, that when this coincidence happens to take place, they express it in terms such as ‘fitting’ and ‘saying the same thing as’, which suggest a highly contingent relation.” (Fernández 2010, p. 79)

We maintain that, taking these reports as a whole instead of quite selectively as Fernandez does, they give better support to our standard proposal. In relation to R6, when the patient says that “... it’s a light feeling where you feel as though I’m actually thinking it ... or you’re receiving it rather ...”, he is saying that he does not have the feeling of being the producer of the thoughts, which suggests a similar interpretation for the part “[I] wonder if

that's me ... it felt like a piece of information." Moreover, the passage "... you find it strange when some different little thought filters through ..." and the related question "why did I think that at this time of day?" suggests that the answer "I don't think that was mine..." concerns who is the producer of a thought occurring in the stream of consciousness at a specific point in time, not who endorses the thought. In relation to R7, the patient saying that "they were being put into me into my mind ... very similar to what I would be like normally" also suggests that the main point of this report is about who is the producer of the thoughts.<sup>17</sup>

Given the standard interpretation, the fact that R6 and R7 say, respectively, that the "inserted" thought "fits with what I suspect" and that her "own thought might say the same thing ..." can be given a completely different interpretation: the patients are saying that they agree with the content of the "inserted" thoughts, although they did not have the feeling of being the producers of these thoughts at the time they occurred in their stream of consciousness.

### **3.3. Broader aspects of the evidence**

Fernández claims that any explanation of thought insertion "should ideally shed some light on both the incorrigibility of delusions as well as their little weight in informing action" (2010, p. 72) and that an advantage of his overall theoretical framework is its consistency with these two facts. Fernández's arguments here are unclear and, if anything, unhelpful to his proposal.

In relation to the first fact, the incorrigibility of delusions, Fernández says:

---

<sup>17</sup> The hesitation and contradiction in these reports concerning who is the producer of the thoughts does not count against the standard approach. It simply reflects the fact that different patients have different levels of confidence about their delusional beliefs of thought insertion.

Schizophrenic delusions, such as delusions of thought insertion, are often very hard to correct by reference to evidence. Thus, it would be hopeless to point out to the patient in case 2 that either God does not exist or He is immortal and, therefore, he cannot possibly be expected to kill God. The patient would not change his mind about the fact that, nonetheless, he believes that he must kill God. (2010, p. 71)

We can see why this would be the case if the commitment hypothesis is correct. (...) Suppose, for instance, that the patient in case 2 experiences being aware of his belief that he must kill God without feeling pressured to endorse the proposition that he must kill God. It is no wonder, then, that no amount of evidence showing that he could not possibly have that obligation is enough to make him drop his belief. (2010, p. 80)

But, in the first passage, Fernandez is conflating the delusional belief that a thought is being inserted with the “inserted” thought itself—it is the delusional belief that is incorrigible in thought insertion *qua* a schizophrenic delusion, not the “inserted” thought. Moreover, the intricate consistency envisaged in the second passage between his theoretical framework and the “fact” of the incorrigibility of the belief *I ought to kill God* is inconsequent, given that no independent evidence showing that the patient has this belief is provided in the first place.

In relation to the fact that delusional beliefs have little weight in informing action, Fernández says:

(...) the commitment hypothesis (...) can shed some light on the fact that delusions of thought insertion often do not make a difference to the patient's behavior. If the patient with thought insertion indeed feels that being aware of some of his beliefs does not commit him to anything about the outside world, then it seems natural that they do not inform his behavior in the way normal beliefs would. Consider one of Bleuler's patients who claims that doctors and nurses are trying to kill him and, nonetheless, eats the food those doctors and nurses give to him. If he did not feel any pressure to endorse the proposition that they are trying to kill him when he is aware of having the belief that they are trying to kill him, then you would expect this patient to take no evasive action. (2010, p. 80)

Again, there is an unwarranted conflation in this passage—he is appealing to a fact about delusions, not even delusions of thought insertion in particular, to make an argument about “inserted” thoughts. Moreover, to explain the fact that the patient does not actually try to kill God for example, it is much more parsimonious, ceteris paribus, to claim simply that the patient does not believe that he ought to kill God or that he disagrees with the command “Kill God.”

#### **4. Conclusion**

We have provided a thorough analysis demonstrating that our proposal dovetails with the evidence whereas Fernández's does not—what is abnormal in the phenomenology of thought insertion is a sense of thought agency in terms of agent causation, not a sense of

thought (or, in particular, belief) commitment in terms of endorsement. However, if there are still fundamental problems with the standard approach, as Fernández would claim, and if these problems also affect our version, the above advantage of our proposal becomes ineffective. We conclude by briefly addressing Fernández's principal criticisms to the standard approach and thereby closing our argument.

Fernández holds that for any version of the standard approach to be defensible it has to accomplish two basic tasks. Firstly, it must identify the experiential features that make a mental event a mental action—i.e. it must provide an explicit characterization of the sense of thought agency. Secondly, it must sustain that the sense of thought agency constitutes the normal experience with the types of thoughts that are considered to be inserted—otherwise, how could one suppose that there is an *abnormal* sense of thought agency in the phenomenology of thought insertion?

Citing leading proponents of the standard approach, Campbell, and Graham and Stephens, Fernández argues that they fail on the first task because they do not provide an explicit characterization of the sense of thought agency in itself—instead, they focus on the background conditions for its presence and on broad analogies with the domain of behavioral action. He grants that Frith, another leading proponent of the standard approach, does provide such a characterization, in terms of the experience of deliberate choice and effort. However, according to Fernández, Frith fails on the second task, since there is no sense of deliberate choice and effort in relation to perceptual beliefs, a type of thought that, he claims, is often an “inserted” thought. This point about the existence of “inserted” thoughts that are perceptual beliefs, he argues, generalizes to any other current characterization of the sense of thought agency—for instance, one in term of awareness of intention.

We accomplished the first task by providing an explicit characterization of the sense of thought agency in terms of the superordinate concept of action, that is, in terms of agent causation, which we distinguished from the sense of agent intentional guidance. In addition, our sense of thought agency differs importantly from current characterizations, which, as Fernández points out, have in mind a notion of the sense of thought agency more in line with the sense of intentional guidance.<sup>18</sup>

Since a fundamental characteristic of the normal experience with thoughts in general is the sense that their occurrence in one's stream of consciousness is produced by the self, intentionally or not, our proposal does not have a problem with the second task. It is also worth noticing that, even though Fernández has not provided evidence that "inserted" thoughts are (perceptual) beliefs of the patient, our approach could handle such types. True, it would be impossible for one to have a sense that a belief is an action, since beliefs are understood as states, and states cannot be understood as actions. However, what is at stake in the phenomenology of thought insertion is the occurrence of a representation in one's stream of consciousness, and the occurrence of factual representations such as beliefs (via memory retrieval, perception or inference) is also normally experienced as being produced by the self.

In sum, our version of the standard approach stands sound, providing the best available answer to the what-question and, hence, circumscribing the explanandum for those who aim to address the why-question.

## References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. (DSM-IV). Washington, DC: American Psychiatric Association.

---

<sup>18</sup> Although, we should point out that some authors in the literature seem to occasionally invoke a notion of sense of agency that is closer to our notion of agent causation—e.g. Campbell, 2002, p. 36.

- Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusions. *Mind & Language*, 20, 163–88.
- Billon, A. (forthcoming). Does consciousness entail subjectivity?: The puzzle of thought insertion. *Philosophical Psychology*.
- Bortolotti, L., & Broome, M. (2009). A role for ownership and authorship in the analysis of thought insertion. *Phenomenology and the Cognitive Sciences*, 8, 205–24.
- Buckareff, A. (2005). How (not) to think about mental action. *Philosophical Explorations*, 8, 83–89.
- Campbell, J. (2002). The ownership of thoughts. *Philosophy, Psychiatry and Psychology*, 9, 35–9.
- Carruthers, G. (2007). A model of the synchronic self. *Consciousness and Cognition*, 16, 533–550.
- Carruthers, P. (2009a). Action-Awareness and the Active Mind. *Philosophical Papers*, 38, 133–156.
- Carruthers, P. (2009b). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32, 121–182.
- Carston, R. (2002). *Thoughts and Utterances: the pragmatics of explicit communication*. Oxford: Blackwell.
- Currie, G. (2000). Imagination, delusion and hallucinations. In M. Coltheart & M. Davies (Eds.), *Pathologies of Belief* (pp.167–182). Oxford: Blackwell.
- Fernández, J. (2003). Privileged access naturalized. *The Philosophical Quarterly*, 53, 252–72.
- Fernández, J. (2010). Thought Insertion and Self-Knowledge. *Mind & Language*, 25, 68–88.
- Fish, F. (1976). *Fish's schizophrenia*, 2nd edn. Bristol: J. Wright.
- Frith, C. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, NJ: Lawrence Erlbaum.
- Gallagher, S. (2000). Self-reference and schizophrenia: A cognitive model of immunity to error through misidentification. In D. Zahavi (Ed.), *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-Experience* (pp. 203–239). Amsterdam: Benjamins.
- Gallagher, S. (2004). Neurocognitive models of schizophrenia: A neurophenomenological critique. *Psychopathology*, 37, 8–19.
- Gallagher, S. (in press). Multiple aspects in the sense of agency. *New Ideas in Psychology*, doi:10.1016/j.newideapsych.2010.03.003

- Garety, P.A., & Hemsley, D.R. (1994). *Delusions: Investigations into the Psychology of Delusional Reasoning*. Oxford: Oxford University Press.
- Gerrans, P. (2001). Authorship and ownership of thoughts. *Philosophy, Psychiatry and Psychology*, 8, 231–237.
- Gertler, B. (2008). Self-Knowledge. *The Stanford Encyclopedia of Philosophy (Winter 2008 Edition)*, Edward N. Zalta (Ed.) URL: <http://plato.stanford.edu/archives/win2008/entries/self-knowledge/>. Accessed October 2010.
- Graham, G., & Stephens, G.L. (2000). *When Self-Consciousness Breaks*. Cambridge, MA: MIT Press.
- Grice, P. (1991). Logic and Conversation. In S. Davis (Ed.), *Pragmatics – A reader* (pp. 305–315). Oxford: Oxford University Press.
- Hoerl, C. (2001). On thought insertion. *Philosophy, Psychiatry and Psychology*, 8, 189–200.
- Jackendoff, R. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
- Jackson, M. C., & Fulford, K.W. M. (1997). Spiritual experience and psychopathology. *Philosophy, Psychiatry and Psychology*, 4, 41–65.
- Jaspers, K. (1963). *General Psychopathology*. Manchester: Manchester University Press.
- Koehler, K. (1979). First rank symptoms of schizophrenia: questions concerning clinical boundaries. *British Journal of Psychiatry*, 134, 236–248.
- Langland-Hassan, P. (2008). Fractured phenomenologies: thought insertion, inner speech, and the puzzle of extraneity. *Mind & Language*, 23, 369–401.
- Mele, A. (2009). Mental action: a case study. In L. O'Brien and M Soteriou (Eds.), *Mental Actions* (pp. 17–37). Oxford: Oxford University Press.
- Mellor, C. S. (1970). First rank symptoms of schizophrenia. *British Journal of Psychiatry*, 117, 15–23.
- Mullins, S., & Spence, S. A. (2003). Re-examining thought insertion. *British Journal of Psychiatry*, 182, 293–298.
- Nichols, S., & Stich, S. (2003). *Mind Reading*. Oxford: Oxford University Press.
- O'Grady, J. C. (1990). The prevalence and diagnostic significance of Schneiderian first-rank symptoms in a random sample of acute psychiatric in-patients. *British Journal of Psychiatry*, 156, 496–500.
- Oltmanns, T.F., & Maher, B.A. (Eds.). (1988). *Delusional Beliefs*. Oxford, UK: Wiley

- Pacherie, E. (2008). The phenomenology of agency: A conceptual framework. *Cognition*, 107 (1), 179–217.
- Peacocke, C. (2007). Mental Action and Self-Awareness. In J. Cohen and B. McLaughlin (Eds.) *Contemporary Debates in the Philosophy of Mind* (pp. 358–376). Oxford: Blackwell.
- Proust, J. (2006). Agency in schizophrenia from a control theory viewpoint. In N. Sebanz and W. Prinz (Eds.), *Disorders of Volition* (pp. 87–118). Cambridge, MA: MIT Press.
- Radden, J. (2010). *On Delusions*. Routledge.
- Strawson, G. (2003): Mental Ballistics or the Involuntariness of Spontaneity. *Proceedings of the Aristotelian Society*, 103/3, 227–257.
- Stephens, G.L., & Graham, G. (1994). Self-consciousness, mental agency, and the clinical psychopathology of thought insertion. *Philosophy, Psychiatry, & Psychology*, 1(1), 1–10.
- Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, 17, 219–239.
- Taylor, M. A., & Heiser, J. F. (1971). Phenomenology: an alternative approach to diagnosis of mental disease. *Comprehensive Psychiatry*, 12, 480–486.
- Thornton, T. (2002). Thought insertion, cognitivism and inner space. *Cognitive Neuropsychiatry*, 7 (3), 237 – 249.
- Vosgerau, G., & Newen, A. (2007). Thoughts, motor actions, and the self. *Mind & Language*, 22, 22 – 43.
- Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.) *Pragmatics* (pp. 230–259). London: Palgrave, London.
- Wing, J.K., Cooper, J.E., & Sartorius, N. (1974). *Measurement and classification of psychiatric symptoms*. Cambridge: Cambridge University Press.
- Young, G. (2008). On how a child's awareness of thinking informs explanations of thought insertion. *Consciousness and Cognition*, 17, 848–862.