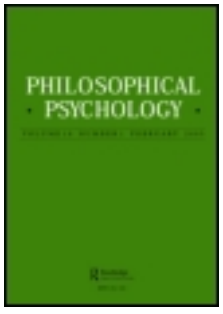

Philosophical Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cphp20

Does consciousness entail subjectivity? The puzzle of thought insertion

Alexandre Billon

Available online: 02 Nov 2011

To cite this article: Alexandre Billon (2011): Does consciousness entail subjectivity? The puzzle of thought insertion, Philosophical Psychology, DOI:10.1080/09515089.2011.625117

To link to this article: http://dx.doi.org/10.1080/09515089.2011.625117

Philosophical Psychology
2011, 1–24, iFirst

Does consciousness entail subjectivity? The puzzle of thought insertion

Alexandre Billon

''There is a thought in me which is not mine.'' This is, roughly, the complaint of patients suffering from thought insertion. This first-rank symptom of schizophrenia is particularly puzzling for it seems to challenge a very well entrenched principle to the effect that our conscious thoughts are necessarily subjective, that we necessarily have a sense of ownership for them (Cartesian principle). Despite their wide disagreement, classical accounts of the symptom save the Cartesian principle by interpreting thought insertion as a problem of the sense of agency for thought rather than as a problem of subjectivity. I argue that those accounts fail and that thought insertion really is a problem of subjectivity. We can nevertheless save the Cartesian principle if we realize that the presupposition, shared by classical accounts, to the effect that inserted thoughts are unequivocally conscious, is ill-grounded. Distinguishing between reflexive awareness and phenomenal consciousness, and relying on a careful comparison between thought insertion and other pathologies of agency, I propose a novel account of the symptom which is compatible with the Cartesian principle and which allows to take the patient's reports seriously. This account, I conclude, opens up novel perspectives on the comprehension of schizophrenia, and reveals a common confusion between two different dimensions of the mind.

Keywords: Agency, Consciousness, Delusions; Immunity Principle; Ownership; Schizophrenia; Self; Subjectivity; Thought Insertion

1. Introduction

Some patients diagnosed with schizophrenia report a strange phenomenon:

> Thoughts are put into my mind like ''Kill God.'' It's just like my mind working, but it isn't. They come from this chap, Chris. They are his thoughts. (quoted in Frith, 1992, p. 66)

Alexandre Billon is Assistant Professor at Univ Lille Nord de France.

Correspondence to: Alexandre Billon, Université Lille-III, Philosophy, Domaine universitaire du ''Pont de Bois'' rue du Barreau, STL, F-59653 Villeneuve d Ascq, France; Univ Lille Nord de France, F-59000 Lille, France. Email: abillon@gmail.com

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/11/000001-24 © 2011 Taylor & Francis
http://dx.doi.org/10.1080/09515089.2011.625117

> 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 onto it like you flash a picture. (quoted in Mellor, 1970, p. 17)

Challenging the cogito, a patient of Hesnard even doubted that he exists by doubting that he owned any thought:

> Am I thinking? Since there is nothing which can prove that I am thinking, I cannot know whether I exist. (quoted in Parnas & Sass, 2001, p. 108)

If we take these reports seriously, some thoughts that seem to the patients to be ''inside them,'' appear not to be ''theirs.'' They seem to be merely inserted in them. Although it occurs in other psychotic conditions, this phenomenon of thought insertion is commonly regarded as characteristic of schizophrenia: along with other disorders of self-consciousness such as alien voices (the patient hears voices), delusions of control (the patient feels as if he is being controlled by some external agency) and thought broadcasting (the patient feels as if his thoughts are publicly accessible), it is a first-rank symptom of this pathology. It is questionable whether this symptom corresponds to a psychologically real phenomenon, discrete from those that underlie other positive symptoms. Yet it is reliably defined and distinguished from neighboring disturbances of self-consciousness (see the recent literature review and conceptual analysis by Mullins & Spence, 2003).

Common description of thought insertion. Inserted thoughts are thoughts the patient describes as being in him but as not being his (and usually, as being somebody else's).

Even though it is quite widespread (it would affect, at one time of another, approximatively one fifth of the patients suffering from schizophrenia), it is still ill-understood.

Inserted thoughts and the immunity principle. Most of the positive symptoms of schizophrenia I have mentioned have something perplexing about them, but thought insertion naturally raises a very specific puzzle, the understanding of which can lead to valuable insights about the nature of consciousness, subjectivity, and maybe schizophrenia in general. The puzzle is the following. On the one hand it seems that we should really take the subjects' reports seriously. Clinicians widely agree that the patients truly mean what they say. Inserted thoughts usually happen only sporadically, sometimes in the midst of normal functioning (typically, when the patients are feeling overwhelmed) and some patients are perfectly rational (at least as far as procedural rationality is concerned). They also explicitly reject metaphorical or watered-down interpretations (Hoffman, 1986, p. 508). On the other hand these reports really sound bizarre. It seems difficult to make sense of them. It is hard, for example, to imagine how we might come to say things like that. This difficulty, emphasized by many philosophers, is now often articulated by saying that the patient seems to misattribute a thought of his and to commit a mistake that we could not possibly make. He would infringe the well entrenched principle to the effect that self-ascriptions grounded on introspection are immune to errors through

misidentification (Campbell, 1998; Gallagher, 2000; Graham & Stephens, 2000). Thought insertion would be puzzling, then, because the patient's reports seem incompatible with such an "immunity principle" (Shoemaker, 1968).

Inserted thoughts and the Cartesian principle. Although I am quite sympathetic to this immunity principle, it is more fruitful to frame the puzzle in terms of a more phenomenologically oriented, and probably more fundamental, principle. This principle does not relate introspection and self-ascription but rather their phenomenological bases, consciousness and subjectivity. After all, introspection is a technical term. It just designates the special, first-personal, access we have to our conscious states. So consciousness is the feature of some states in virtue of which they can be *self-ascribed* based on introspection. In the same way, if we call subjectivity the way my conscious states typically appear to me to be *mine* "from the inside," in the first-person, then it is in virtue of their subjectivity that I self-ascribe some states on introspective grounds. In that sense the immunity principle seems to be derivative of a principle to the effect that consciousness entails subjectivity. This principle has received many formulations in the history of philosophy. Some thinkers of the phenomenological tradition, who characterize subjectivity as a form of self-awareness say that all awareness is self-awareness. Jaspers, for example, claimed that "we are not able to have any clear sight of any psychic event without our self-awareness being involved" (1913/1997, p. 578), and he indeed believed that it made thought insertion strictly impossible to understand. James said that "every thought is part of a personal consciousness" (1890/1983, p. 225) and he subsequently made clear that this was not meant to apply to subconscious thought. More recently, Kriegel (2004) characterized subjectivity as a "peripheral" form of intransitive self-consciousness and argued that it is entailed by consciousness. But the first and foremost modern exponent of this principle is probably Descartes (1641/1985). Descartes called "thoughts" what we would call "conscious thoughts" and he believed that the cogito established the certainty not only that "there is a thought," as Lichtenberg (1806/1990) would have it, but also, that this thought seems and cannot fail to be mine.[1] I will for that reason, call it the "Cartesian principle."

> **Cartesian principle**. If a thought is conscious then it is subjective (in the sense that it appears to me to be *mine* "from the inside").

The puzzle of thought insertion can now be reframed as the apparent incompatibility of the patient's report with the Cartesian principle.

Outline. The aim of this paper is to address this puzzle. Here is how I will proceed. In the first three sections I will present the way classical accounts of thought insertion answer this puzzle. These accounts fall roughly into two separate categories. Some researchers claim that in spite of appearances, thought insertion is in fact compatible with the Cartesian principle. I will argue that those compatibilist accounts fail (section 1). Others argue that thought insertion requires us to dismiss the Cartesian principle. I will argue that these incompatibilist alternatives are not better off

(section 2). All such accounts share a few common assumptions that seem unwarranted. In particular, they all interpret thought insertion as a problem of agency and they all consider that inserted thoughts are unequivocally conscious (section 3). I will reject these assumptions and present my own compatibilist account. In a nutshell, inserted thoughts really lack subjectivity but they are not conscious either, at least in the usual sense of the term. The subject certainly has a good reflexive awareness of those thoughts but it is not clear that they are phenomenally conscious. And while phenomenality entails subjectivity, it is not clear that such a form of good reflexive awareness should entail more than a spatial analogue of subjectivity, namely, apparent inner location. I claim that inserted thoughts are thoughts, or thought-like processes, which the patient disowns because they are phenomenally unconscious, even though he has a good reflexive awareness of them. This allows the account both to save the Cartesian principle and to respect the phenomenology of inserted thoughts.

2. Compatibilist Accounts

Thought insertion, as previously mentioned, threatens the Cartesian principle. This might not actually be as straightforward as it sounds. We took for granted, for example, that the patients' reports are perfectly meaningful and that they truthfully reflect their phenomenology. But there may be reasons not to make these assumptions.

2.1. Denying that the Patients' Reports Are Intelligible

Coliva (2002) rightly contends that if we agree that the Cartesian principle stems from an obvious a priori truth connecting introspection or consciousness and subjectivity, then we cannot take the patients' reports at face value.[2] Doing so would be self-defeating: it would make their accounts clearly contradictory and jeopardize their very meaningfulness.

This rescue of the Cartesian principle might appear like an ad hoc maneuver. After all, this principle is what was at stake in the first place, and we have seen that some patients who seem otherwise rational report inserted thoughts. Coliva finds independent reasons to question the rationality of the patients and the meaningfulness of their reports in the fact that they are considered mentally ill:

> Why should we consider her report as an expression of some kind of cognitive illusion, which we take as a symptom of mental illness, as opposed to, at most, a possible mistake in identifying the producer of the thought? And, connectedly, why should we try to cure her, rather than just, at most, correct her? (2002, p. 42)

There are however many acceptable criteria for mental illness and many reasons to treat someone which do not involve any rationality problem. We should treat patients mainly because they suffer and complain of their condition, and we should

treat them even in situations where there is no need to correct them. Mental illness does not entail meaninglessness.

More generally, it is only because we assume that patients are ''rational'' or at least sufficiently similar to us so as to be intelligible and to mean what they say that thought insertion can be deemed puzzling in the first place. Prima facie I will assume that this impression of puzzlement is right and I will suppose that unless it has proven impossible to do so, we should try to save the intelligibility of the patients.

2.2. Distinguishing Ownership from Agency

Some compatibilist accounts do acknowledge the intelligibility of the patients' reports. In order to argue that, even so, thought insertion does not threaten the Cartesian principle, these accounts rely on the distinction between agency or authorship and ownership. When I move my arm, I not only feel that my arm is moving, I also feel that I am the one moving my arm. In other words, I not only have a sense of ownership of the movement, associated with the subjectivity of the experience of my moving body, I also have a sense of agency for it. Compatibilists believe that thinking is a kind of action, so that the agency-ownership distinction applies to thoughts as well as to movements of limbs. Compatibilists further argue that delusions of thought insertion involve a misidentification of thought-agency rather than a misidentification of thought-ownership, leaving the immunity principle unchallenged. The patient's thought would be his own, it would be subjective and he would implicitly admit it. He would only fail to acknowledge that it was generated by his own agency (Gallagher, 2000; Graham & Stephens, 2000, p. 121).

Philosophers who appeal to the agency-ownership in the explanation of thought insertion were highly influenced by the works of Chris Frith (Frith, 1992; Frith & Done, 1988) from which they draw important support (see Blakemore, Wolpert, & Frith, 2002; Frith, 2005, for recent developments of the model), Frith has posited that the control mechanism which predicts the sensory outcome of our actions (the so-called ''forward model'') and compares it to the actual sensory outcome could also be used to distinguish sensory changes that are self-generated from sensory changes that are not. This mechanism would be impaired in schizophrenia, which would explain patients may suffer from delusions of alien control. Frith (1992) extended this theory to account for other passivity phenomena that concern thoughts rather than bodily actions. The basic idea is that thought is also a motor process, so thought insertion can be explained by the same kind of prediction problem. A subject would produce a motor command for a thought, but, failing to predict the accurate outcome of that motor process, he would fail to control the thought and to self-attribute it. It is not, however, totally clear how we should understand the sense of control and agency for thoughts as opposed to bodily movements. Campbell (1998, 1999), Gallagher (2000), and Graham and Stephens (2000) all argue that it is accordingly not clear how we should understand Frith's proposal at the personal level, and they try to amend his explanation in order to deal with this problem. Compatibilist accounts like those of Gallagher (2000) and of Graham and Stephens

(2000) try to do that while maintaining the immunity principle and the Cartesian principle. They retain Frith's idea that thought insertion would involve a problem of agency rather than subjectivity.

The main problem facing such interpretations consists in accounting for the difference between thought insertion and other phenomena affecting the sense of agency for thought, like (i) controlled thoughts, and (ii) intrusive thoughts. We will discuss the two phenomena with particular attention to the second one, as it has been extensively studied over the last thirty years. Psychiatrists usually distinguish inserted thoughts, which are thoughts the patient explicitly disowns, and controlled thoughts, which the patient regards as simply under the control of someone else's agency (Jaspers, 1913/1997, pp. 122–123) opposes "implanted thoughts" to "passivity thinking," while O'Grady (1990) opposes "inserted thoughts in the narrow sense" to "inserted thoughts in the wide sense." Patients seem to make a distinction between the two phenomena; they can claim ownership for controlled thoughts (Mullins & Spence, 2003), and it has been shown that controlled thoughts are less specific to schizophrenia (O'Grady, 1990). Yet it seems difficult to account for the difference between the two phenomena without claiming that, whereas controlled thoughts result from a deficit in the sense of agency, thought insertion stems from a disruption of the sense of ownership. This is not a possibility which is open to those who claim that the sense of ownership is intact in schizophrenia.

Consider intrusive thoughts now. We all commonly experience passive, unbidden thoughts entering our stream of consciousness. We seem to have no sense of agency for these thoughts. Yet we do not disown such thoughts, nor, a fortiori, do we attribute them to someone else. The difference between unbidden and inserted thoughts is not merely a question of degree. Patients diagnosed with obsessive compulsive disorders (OCD) are incessantly haunted with intrusive thoughts that they constantly try to resist. These thoughts are often very unpleasant, and the patients' vain attempts to suppress them ironically increase their recurrence (try not to think about a red car, and you will experience this so-called "red car effect"). At some point, the occurrence of the thought will seem literally irresistible and the associated feeling of passivity will be radical (see e.g., Purdon, Rowa, & Antony, 2005; Salkovskis, 1989). Even then, patients with OCD will acknowledge that the intrusive thoughts are really theirs.

Thought insertion, controlled thinking, and obsessional thinking are different symptoms, but is seems hard to account for their differences once it has been claimed that thought insertion is actually a problem of the sense of agency. Confronted with this problem, the compatibilist must expand his account of thought insertion by spelling out what is peculiar to the alien feeling of inserted thoughts. Despite their differences, all the accounts I am aware of end up citing some kind of psychological discrepancy as an explanation for the alien feeling. Graham and Stephens (2000) cite a discrepancy between the content of inserted thoughts and the patient's self-view. Gallagher (2000) cites a discrepancy in the diachronic phenomenal consciousness of the patients. I tackle these explanations in turn to show why they fail.

Graham and Stephens's (2000) explanation is top-down. The content of the inserted thought would not fit the patient's narrative, which would explain why the patient is prone to attribute them to somebody else.

> Suppose that Mary, a young mother concerned with her child's welfare and her own maternal responsibilities, finds the thought "Bad mother" or "You're hurting your child" or "Joan Crawford!" occurring in her stream of consciousness. She does not acknowledge in herself or attribute to herself the sorts of intentional states that would naturally find expression in these thoughts. She pictures herself as a caring, competent mother . . . . Thus, she may have the strong impression that someone is speaking in, or speaking to her. (Graham & Stephens, 2000, p. 173)

By contrast, Graham and Stephens (2000) claim that the intrusive thoughts of OCD patients have a less specific content. They would fit the patient's self-view, and this would explain why their thoughts do not feel alien.

Such an explanation should be rejected because it relies on a description of intrusive thoughts which is inaccurate. It is true enough that patients with OCD self-attribute their intrusive thoughts and consider them as representative of their personality (Salkovskis, 1985). This is indeed a central element of the disorder. But the self-attribution cannot be explained by the narrative fit. On the contrary, it is because the patient (rightly) self-attributes the intrusive thought that he fears it might be representative of his personality. This is shown by studies of nonpathological intrusive thoughts. 90% of the population has experienced such (benign) intrusive thoughts. These thoughts have similar content to pathological intrusive thoughts (Rachman & da Silva, 1978). Without the belief that intrusive thoughts reveal something deep about one's personality, they are easily dismissed and their frequency does not become pathological (Salkovskis, 1985). In the same way, patients recovering from OCD still experience the same intrusive thoughts, still self-attribute them, but no longer believe that these thoughts are in any way representative of their personality (Freeston & Ladouceur, 1999). Accordingly, cognitive models of OCD explain the patient's self-view by the self-attribution of the thought rather than the other way around:

> Individuals prone to developing Obsessive Compulsive Disorders (OCD) may have difficulty assimilating obsessional thoughts into their existing self-view and instead worry that their view of themselves is inaccurate. For example, the loving parent who has a thought of harming her child can assimilate the thought (e.g., "even a good, loving parent like me can have a thought like this"), or begin to accommodate the thought (e.g., "maybe I'm a homicidal maniac"). In the latter case, the occurrence of the thought itself is the only piece of evidence that undesirable personality characteristics are lurking [emphasis added]. As such, its absence would signify the absence of these undesirable characteristics so the individual becomes highly invested in not having the thought and is exquisitely sensitive to its recurrence. (Purdon, 2004, p. 123)

Gallagher's (2000) account is more bottom-up. Instead of explaining the alien feeling in terms of a discrepancy in "narrative consciousness," he views it as

stemming from a discrepancy in diachronic phenomenal consciousness. Following Husserl, Gallagher claims that we should posit some kind of central anticipation mechanism in order to explain various experiences tied to expectations; like, for example, the experience of surprise. Non-psychotic thoughts would be prefaced by anticipation or, as Gallagher puts it, a ''protention,'' but inserted thoughts would not, and this would explain the difference in phenomenology. It is not easy to see how this is supposed to differentiate inserted thoughts from other pathologies of thought agency. To explain why unbidden thoughts do not feel inserted, Gallagher seems committed to the view that such unbidden thoughts cannot happen unexpectedly, by surprise, or in a discontinuous manner. ''Protention,'' claims Gallagher, ''provides some kind of expectancy for [unbidden thoughts], even if it is completely indeterminate'' (Gallagher, 2000). But unless one assumes that nothing but inserted thoughts can conceivably count as really surprising and ''not even indeterminately expected'' thoughts, I do not see why we should accept such claims. Gallagher's (2000) account, then, would have to beg the question.

2.3. A Methodological Excursion

There is a potential line of reply for those who consider thought insertion to be a problem of agency. In order to distinguish thought insertion from controlled thoughts and obsessive thinking, we had to take the patients' reports quite seriously. We considered that those reports were not only meaningful but that they reliably reflected their phenomenology.

This supposition can explain why thought insertion is puzzling. Thought insertion should not be considered puzzling, as we said, if the patient's reports are not meaningful. But the same would go if they were, for example, much worse at introspecting their phenomenology than we are. In such conditions, their reports would be no threat at all to the Cartesian principle. The puzzle arises because (i) we are reliable witnesses of the subjectivity of present occurrent thoughts, and because (ii) we suppose that the patients are similarly reliable:

- Whether a given thought, which is occurring now, is subjective is normally a transparent aspect of our phenomenology: it is something we can normally know simply by introspecting our experience. (It is not, say, like telling whether that wine tastes rather like a Bourgogne or like a Bordeaux. Notice also that unlike the subjectivity of an occurrent thought, its ''felt owner'' is not normally an aspect of our phenomenology: I can tell by introspection whether a thought is subjectively mine or not, not in general whether it is subjectively mine, John's, or Chris's).
- Our reports about subjectivity reliably witness this aspect of our phenomenology.
- We suppose that the patients' introspective abilities are not worse, or not much worse than ours, in that respect.
- We suppose that their reports are meaningful and reliably witness the subjectivity of their thoughts as well.

There is no independent evidence that the patients are not relevantly similar to us, no independent evidence, that is, indicating that we should consider our

introspective reports and the patient's testimonies with different standards. To the contrary:

> - As already mentioned, the usual reports characteristic of thought insertion (''this thought is in me but it is not mine, it is his'') have been reliably defined and isolated.
> - Some patients are no less rational than we are and reject metaphorical interpretations of their reports. (Hoffman, 1986, p. 508)
> - Patients suffering from passivity symptoms have a theory of mind (the ability to think about thoughts), and even though they often lack insight, it has been shown that their self-reports are generally reliable. Even if they might fail to attribute them to illness, they are, for example, generally aware of their symptoms and personality traits. (Bell, Fiszdon, Richardson, Lysaker, & Bryson, 2007)

In these conditions, I will suppose that we should, by default, consider that the patients as relevantly similar to us. I will suppose, more specifically, that, for lack of independent evidence to the contrary, we should by default stick to the following phenomenological constraint: inasmuch as they concern such transparent aspects of their phenomenology like subjectivity, the patient's reports should be taken at face-value. In the case at hand, we should by default admit the following:

> **Phenomenological constraint.** If the patient says that an occurrent thought is not his, then it is not subjective. If the patient says that an occurrent thought feels his, then it is subjective.

3. Incompatibilist Accounts

These compatibilist accounts owe their failure to the fact that they interpret thought insertion as a problem of agency rather than ownership. The desire to salvage the immunity or the Cartesian principle often motivates their focus on agency. But maybe we should consider giving up those principles, put forward an incompatibilist account of thought insertion, and develop a ''complex view'' of subjectivity according to which it does not boil down to consciousness. In a series of influential papers, the Campbell (1998, 1999, 2004) has done just this. Interpreting thought insertion as a problem of ownership, he suggests that it challenges the immunity principle. More precisely, he claims that this principle is not true when it comes to occurrent thoughts which express long standing dispositional states such as beliefs and desires (Campbell, 1998, p. 110, 1999). Accordingly, consciousness would not in general be sufficient for subjectivity. What additional ingredient is required? According to Campbell, there are two strands in our notions of ownership and subjectivity. A thought will seem fully mine both (i) if it is conscious (or equivalently, again, ''accessible through introspection'') and (ii) if it seems to be caused by me, that is, if I have a sense of agency or authorship for the thought (Campbell, 1999, 2004). Campbell further analyzes this agency for a thought as ''the fact that [it is] the product of my long-standing beliefs and desires'' (1999, p. 621). In normal thinking, my background beliefs and desires would cause a subpersonal motor instruction for a thought to be issued, which would cause both the occurrence of the thought and of

an efferent copy. The comparison between this efferent copy (which, Campbell makes it explicit, would not correspond to a sense of effort) and the thought would result in the sense of agency. A problem with this comparison mechanism would account for thought insertion.

There is something quite disappointing about such an account, for it turns out to be very close to those we have already scrutinized. After all, it also pictures thought insertion as a problem of agency. The main difference is the following: whereas Campbell considers that agency is required for full-blown subjectivity, others consider that it is independent of subjectivity. Campbell's account will thus be subject to the same kinds of objections. In the same way as Gallagher (2000) and Graham and Stephens (2000), Campbell cannot at the same time take the patient's report about his phenomenology seriously (and in particular, meet the phenomenological constraint) and account for the difference between thought insertion and obsessions or merely controlled thoughts. By default, we should accordingly reject his theory.

4. Common Assumptions

Why do all of these explanations fail? Despite their disagreements, all of them share a surprising number of presuppositions:

1. They interpret thought insertion as a problem of agency. Incompatibilists, like Campbell, simply claim that agency is required for full-blown subjectivity.
2. Their explanation of this problem cites a psychological discrepancy.
3. They claim that inserted thoughts are still subjective in some sense. Campbell (1999, 2004) admits that among the two strands of our notion of subjectivity, one, which is a priori tied to consciousness, applies to inserted thoughts.
4. They suppose that inserted thoughts are accessible through introspection, and thus conscious.

These presuppositions are not unrelated. Supposing that the inserted thought is conscious, philosophers and psychologists assume that it must be subjective in some sense because they stick to a version of the Cartesian principle, albeit a weak one, which connects a priori consciousness and accessibility through introspection with subjectivity. But once this step has been completed, the margins of interpretation are severely reduced. The patient says that the thought is not his. If he cannot be denying that the thought is subjective in some sense, then, one could reason, he must be denying that he has a sense of agency for it! The claim that thought insertion results from a problem of agency can further be supported by the fact that, following Frith, many psychologists consider that most positive symptoms in schizophrenia can be explained by deficits in agency. Finally, as I have agency for an action only if it is caused in the appropriate way by an intention of mine, that is, only if a certain causal relation holds between my action and other psychological states, it should come as no surprise that whoever explains thought insertion in terms of agency will explain it in terms of a discrepancy between the thought and other psychological states.

However motivated they are, these common presuppositions are wrong. Consider the first one. We already saw that assuming that thought insertion stems from a

problem of agency does not seem to leave enough logical space to discriminate it from other acknowledged pathologies of agency. Some might think that we should nevertheless adopt an account in terms of agency in order to comply with the vast majority of studies which follow Frith in explaining most schizophrenic symptoms as problems of agency. We can, however, perfectly well accommodate these studies by explaining thought insertion in terms of a deficit which is logically prior to deficits in agency and which can therefore account for those deficits. A general disorder in the sense of ownership could, for example, adequately explain deficits in the sense of agency. Indeed, the sense of agency is arguably nothing but the sense of owning the intention which appropriately causes and controls the relevant action.

More generally, it seems that I could still perceive a given thought as mine, whatever its relations to other mental states. Even if I felt that it was generated and controlled by someone else, even if I felt that it did not fit my self-view, I could still perceive it as mine. Conversely, no relation to other mental states seems to account for the fact that a given state does not appear mine to me. We can say that subjectivity is seemingly or phenomenologically intrinsic: it seems that, whatever its relations to other mental states, a given mental state could have or lack subjectivity. It is accordingly very hard to understand how it could be explained by a property, like psychological discrepancy, that is not phenomenologically intrinsic itself.

4.1. Does Apparent Inner Location Entail Subjectivity?

If both compatibilists and incompatibilists assume that thought insertion is a problem of agency, it is because they presuppose that inserted thoughts are still subjective, that the patients have a sense of ownership of them. But this presupposition seems ungrounded in the first place. Patients do not say that, although they have no agency for the thought, it nevertheless seems theirs. They explicitly deny this! Interpreting thought insertion as a problem of agency rather than ownership directly conflicts with the requirement to take the patient's reports about the transparent aspects of his phenomenology at face value (phenomenological constraint). It obliges us to reinterpret their reports. Not only that, it obliges us to reinterpret their reports in a very awkward way, for who would say "this is not x, but y" to express (metaphorically or otherwise) "this is *x*, but not *z*"? Theorists are well aware of this problem, and they usually lay down an argument to the effect that a reinterpretation (but why, one might further ask, such a weird one anyway?) is independently needed. Strangely enough, the argument does not rely on independent knowledge about the patient's cognitive functioning. It relies on the very report that seems to contradict its conclusion. It is usually left rather implicit, but it can be construed as follows.

1. Patients say honestly that the thought is "in them, but not theirs."
2. If a patient honestly says that a thought is "in him," then he judges that the thought is in him.
3. If a patient judges that the thought is in him, then that thought seems to him to be in him ("apparent inner location"), and he is reflexively aware that the thought is in him ("reflected-on inner location").

4. **INNER-LOC (SUB).** If a thought seems to someone to be in him (apparent inner location), or at least if that someone is also reflexively aware that the thought is in him (reflected-on inner-location), it seems to him to be his (subjectivity).

(1) and (2) are trivial. (3) can be granted. This premise distinguishes, it should be noted, the explicit reflexive awareness which underlies the judgment of inner location (reflected-on inner location) from the more basic apparent inner location on which it is based. Apparent inner location is to reflected-on inner location what subjectivity (apparent ownership) is to reflected-on subjectivity. Unless otherwise noted, I will always use 'reflexive awareness' to designate explicit forms of reflexive awareness.

The fourth premise is the decisive one. True enough, it might be correct in usual circumstances that apparent inner location and reflected-on inner location come with subjectivity. But we need more than that, for we have seen that there are reasons to claim then even though it is normally true, it is not true in the particular case under consideration. Remember: the claim that the inserted thoughts are subjective, implied by (4), not only has us infringe the phenomenological constraint. More broadly and more simply here, it has us contradict a perfectly symmetrical argument which would conclude that inserted thoughts are not subjective from the fact that the patients report that the thoughts are not theirs, using only premises that are more obvious and have a better claim on generality than (4). So those who claim that inserted thoughts lack subjectivity need to show that (4) is always true, and even maybe, obviously so. They usually try to do that by claiming that it is an analytic truth. This would imply that the patient actually means that the thought is his when he says that it is in him. That would indeed make it charitable to claim that the patient, if he is to avoid obvious inconsistencies, does not mean to disown the thought when he says that the thought is not his.

I take it that most of the philosophers I criticized here endorse something like the analytic truth of INNER-LOC (SUB). They usually commit to it by successive, "sorite" shifts in meaning. They start by saying that the thought "seems to be in the patient" and with the help of a few spatial metaphors for ownership ("within his ego boundaries," "within his stream of consciousness," "within his inner space") they end up saying that it seems to be his (Campbell, 1998, p. 109; Gallagher, 2000; Graham & Stephens, 2000; Synofzika, Vosgeraub, & Newen, 2008). INNER-LOC (SUB) is not however, analytically true. What does it mean to say that a thought seems to be in me? "In me" literally has a spatial meaning. It could mean "under my skin and skull" or more broadly, and more plausibly, "that is always where I am, that follows me around." Many philosophers have insisted on the fact that thoughts have no spatial phenomenology. This might be a reason to interpret the talk of apparent inner location in terms of subjectivity. The point about the phenomenology, however, is misguided. True enough, as noticed by Descartes (1641/1985, p. 59) and Hume (1740/1978, p. 236), from the inside, my thoughts do not seem to be extended in me. But this does not imply that they have no spatial phenomenology. For even if my conscious thoughts do not seem to be extended in me, almost all of them seem to be located in me. One can easily account for this apparent inner location by the following phenomenological fact: it seems both that (i) I do not have to "move

around'' to be aware of them, and that (ii) they ''follow me around.'' As long as my conscious thoughts exist, moving my head or walking away is not sufficient to avoid them. In other words, my awareness of my conscious thoughts generally seems both quite easy (there is no effort I need to make in order to access them) and robust (this availability is insensitive to my movements). By contrast, neither perceptual objects like tables and chairs, nor my own unconscious mental states, nor other people's experience satisfy (i) or (ii). They seem to be outside my ego boundaries. Such a seemingly easy and robust availability to reflexive awareness, or for short, such a good availability to reflexive awareness, precisely captures the sense in which my conscious thoughts normally seem to be in me. Accordingly:

- Apparent inner location is a good availability to reflexive awareness.
- Reflected-on inner location is a good form of reflexive awareness.

Now we can ask: do reflexive awareness of that sort and the good availability on which it is based entail subjectivity? The answer is no. Subjectivity is the ''meishness'' typical of my conscious states, the property in virtue of which they always appear mine to me. It has been described as an implicit kind of self-awareness (Kriegel, 2003) or self-affection (Henry, 1990), and it is plausible that subjectivity and reflected-on subjectivity entail both a good availability to reflexive awareness and a good form of reflexive awareness. But there are good reasons to think that the converse does not hold.

First, it is easy to imagine cases, and even, we shall see, empirically plausible cases, in which good reflexive awareness comes without subjectivity. I can imagine having a small inner monitoring system (a kind of ''inner eye'') constantly informing me about my unconscious states or about your mental states. I could have, to quote Mellor's (1970) patient, ''a screen'' in the middle of my skull where you could ''flash'' pictures conveying your thoughts. This would be a case of good reflexive awareness of thoughts that do not seem mine. Take a more mundane example: there are sounds that we actually hear inside us, like tinnitus, but that do not seem to be the vehicle of our mental states. Nonetheless, I can imagine a situation in which they express some mental states. I suspect, finally, that we all know empirical cases of such dissociations between good reflexive awareness and subjectivity: hearing voices is having a good ''reflexive'' awareness of things that are not subjective mental states (they do not have the meishness typical of our conscious states). In all such examples, there is a thought or a thought-like process that has both apparent inner-location and reflected-on inner location without seeming mine the way subjective thoughts do.

Second, and more deeply, there is an important modal difference between apparent and reflected-on inner location on the one hand and subjectivity on the other. Whereas it seems almost impossible to imagine of one of your subjective states that it is actually not yours, in all the hypothetical cases just considered, I could definitely imagine that the thought or thought like process I am aware of, and which seems in me, is not actually mine, but that by some neurosurgical feat, I am being linked to someone else whose thoughts are being communicated to me.

4.2. Does Apparent Inner Location Entail Phenomenal Consciousness?

Let us take stock. The puzzle of thought insertion comes from the apparent contradiction between the common description of thought insertion (inserted thoughts are thoughts the patient describes as being in him but as not being his) and the Cartesian principle (if a thought is conscious then it is subjective). We have seen actually that the common description can only threaten the Cartesian if we take the patients' reports seriously, as the phenomenological constraint recommends (if a subject says that a thought is not his, then it is not subjective). Theorists thus seem to be faced with a choice between the Cartesian principle and the phenomenological constraint. Traditional compatibilists must reject the phenomenological constraint and assert that inserted thoughts are actually subjective. Incompatibilists officially reject the Cartesian principle (or, again, one of its siblings, like the immunity principle), although we saw that, at least implicitly, they maintain some version of it and infringe the phenomenological constraint. Finally, we saw that even if patients report the inserted thought as being in them—and that they arguably seem to be located in them—this does not provide an independent reason to rebut the phenomenological constraint and affirm that the patient's inserted thoughts are subjective. In our present state of knowledge, the only option available in order to maintain the common presupposition to the effect that inserted thoughts are still subjective in some sense is thus to argue, by modus tollens, for the falsity of the phenomenological constraint from the certainty of the Cartesian principle. But would such an argument really be cogent?

It is now time to question the important supposition on which the claim that the Cartesian principle is inconsistent with the phenomenological constraint and the common description of thought insertion—and the puzzle of thought insertion—has been relying. We have been supposing all along that inserted thoughts were conscious, which combined with the Cartesian principle entailed that they are subjective. In a sense this seems obvious. The patient reports them easily, just like one would report conscious thoughts. Consciousness, however, is a famously ambiguous and slippery notion (e.g., Block, 1995; Chalmers, 1996; Siewert, 1998). As Block puts it, "there are a number of very different 'consciousnesses'. . . . These concepts are often partly or totally conflated, with bad results" (1995, p. 227).

What "consciousness" do the patient's reports license? The patient reports his thoughts, so he is reflexively aware of them. He reports them as being "in him," which shows, as we have seen, that his reflexive awareness is good. But at least prima facie, his reports do not support more than that. Now, to what kind of "conscious" thoughts does the Cartesian principle attribute subjectivity? The Cartesian principle, I take it, requires consciousness in the phenomenal sense. My phenomenally conscious states are those states there is something it is like to be in, and it is because what it is like to be in such states is always what it is like for me to be in them that they are subjective. Whereas reflexive awareness is tied to reportability and measures a quality of informational access, phenomenal consciousness is tied to qualitative character and affectivity. It should be noted however, that even if affects and

emotions are paradigmatic of phenomenal states, perceptions, episodes of inner speech, and even thoughts are phenomenally conscious. There is something it is like to think that P: there is a cognitive phenomenology, or so will I admit.[3]

So here we are: the Cartesian principle is only warranted for phenomenally conscious thoughts, but at least prima facie, the description of inserted thoughts only warrants that the patient has a good reflexive awareness of them. So the standard description of inserted thought is not as such in conflict with the Cartesian principle and the phenomenological constraint. Such a conflict—and the puzzle it gives rise to—only arises if it is assumed that good forms of reflexive awareness always entail phenomenal consciousness, that is, if it is assumed that the following INNER-LOC (PHEN) is always true:

- **INNER-LOC (PHEN).** If a thought seems to someone to be in him (apparent inner location), or at least if that someone is also reflexively aware that the thought is in him (reflected-on inner-location), it is phenomenally conscious.

Otherwise it would be easy to claim that inserted thoughts, although ''reflexively conscious,'' are neither phenomenally conscious nor subjective. But what reasons do we have to believe that INNER-LOC (PHEN) is always true? We have no reason to think that it is a priori true or empirically true. I tackle those two points in turn.

The arguments introduced in the preceding subsection to show that apparent and reflected-on inner location do not entail subjectivity can be extended to show that they do not a priori entail subjectivity. Just replace ''subjectivity'' by ''phenomenal consciousness'' in that previous subsection. There does not seem to be any nomological connection between the two either. The empirical study of consciousness does not, in the present state, indicate the existence of such a link.[4] We can accordingly rightfully doubt that INNER-LOC (PHEN) must always be true.

One might be tempted to argue that INNER-LOC (PHEN), and actually INNER-LOC (SUB), are always true by appealing to so-called ''monitoring'' theories of consciousness and subjectivity (Kriegel, 2006). Partisans of these theories try to explain consciousness and subjectivity—they take the latter to be essential to the former—as stemming from a kind of reflexive awareness. A state of a subject would be conscious and subjective in virtue of being reflexively represented by the subject. Of course, not just any kind of reflexive awareness will do: there are plenty of unconscious things, and plenty of thoughts of which I am aware but which are neither conscious nor subjective (the trees outside, my desire to marry my mother, etc.). Only the forms of reflexive awareness that seem immediate might do. According to Rosenthal (1997), reflexive forms of awareness which do not rely on conscious inferences or observation will be appropriately immediate, and will entail consciousness and subjectivity.[5] Others consider that this is insufficient and argue that only a self-reflexive awareness will be appropriately immediate. According to these ''same order'' theorists, a conscious and subjective state must actually be represented by itself rather than by a distinct higher-order state (Kriegel, 2003). I am not convinced by either view, but if one of them were true, I would argue that reflexive awareness can be good (seemingly easy and robust) without being of the

form (appropriately immediate) that might explain consciousness and subjectivity according to this view. To a partisan of monitoring theories the arguments we gave can indeed be seen as arguments for the existence of forms of reflexive awareness which are not appropriately immediate even though they are good.

5. A Solution in Terms of Phenomenal Consciousness

Once the presuppositions shared by common accounts of the symptom are abandoned, it is very easy to form an hypothesis that can solve the puzzle or thought insertion. According to that hypothesis, INNER-LOC (SUB) and INNER-LOC (PHEN) would be false in the case of inserted thoughts, so there would be no conflict between their common description, the Cartesian principle and the phenomenological constraint. The patient would have a good reflexive awareness of their inserted thoughts but those would lack phenomenality and subjectivity altogether.

Pace Jaspers (1913/1997), the distinction between different forms of consciousness would thus allow us to understand the phenomenology of those enigmatic ''made'' symptoms. Inserted thoughts would be thoughts or thought-like processes of which the patient is reflexively aware although they are not phenomenally conscious for him. They would be vehicles of thought which seem to be in his mind but lack phenomenal consciousness. This does not mean that there is no phenomenology associated with thought insertion, only that it is, so to speak, an extrinsic, or a ''second-order phenomenology'': the phenomenology of being presented, in a way that is phenomenally conscious, with thought processes in one's mind that are not, in themselves, phenomenally conscious. Inserted thoughts would then have a kind of Janus-faced profile:

- They would differ from the patient's ordinary thoughts by being phenomenally unconscious for him. They would, in that respect, be somehow akin to sentences, images, unconscious computational processes, and other people's conscious thoughts.
- But they would differ from such sentences, images, unconscious computational processes, and other people's conscious thoughts in that they are ''apparently in the patient'': he has a good reflexive awareness of them.

Thought insertion would thus fundamentally be a disorder of phenomenal consciousness. More precisely, it would be a differential disorder: too much good reflexive awareness as compared to phenomenal consciousness.

This hypothesis has many explanatory virtues. It is, first of all, compatibilist and thus conservative: it is consistent with the Cartesian principle. It also complies with the phenomenological constraint. To see why, compare what it is like to have a good access to a thought that is phenomenally conscious for you and to a thought or a thought vehicle that is not. When you attend carefully to one of your phenomenally conscious states, say an experience of a red apple, you are immediately aware of the features of its content. The red apple appears through the experience in a *perfectly* transparent maner. It would be wrong, however, to infer from this that your experience is *wholly* transparent, for you are also immediately aware of something

which seems to be an intrinsic feature of your experience itself, namely, its being an experience of yours, its subjectivity. You are not only aware of a red apple, but also, that you are having an experience of it. This feature is so trivial and ubiquitous that it is hardly noticed, but it is there. Experiences are, so to speak, like glasses with their frames: only a part of them is transparent, even if this part is indeed perfectly transparent.[6]

We cannot directly compare this situation to one in which you have a *good access* to a phenomenally unconscious thought—something we normally never experience—but we can compare it to something near enough. Consider a case in which you have a not-so-bad access to a thought through a phenomenally unconscious vehicle. You are, for example, attending to a picture of the same red apple. If your access is very bad (say you have forgotten your glasses) you might have to interpret the image and fail to be immediately aware of what it represents. If your access gets better though (you just found your glasses), the vehicle might let the red apple appear more transparently, like in the conscious case. Like in the conscious case too, however, it will not become totally transparent. For when you attend to the thought vehicle, you won't just be aware of the red apple, you will also be directly aware of that this image representing a red apple is not a conscious state of yours: you will be aware of its lack of subjectivity. We could take other examples and compare, say, one of your conscious thoughts about an apple and the voice of someone you hear, who utters something with the same content. Each time, the phenomenology of the two situations differ by the subjectivity of the thought vehicle you are attending to. It differs also, by the apparent inner location (in those examples the unconscious thought and its vehicle seem outside you, your access is not that good), but I have argued that this last difference cannot explain the difference in subjectivity. It should be explained, then, by the difference in phenomenality.

Among the mental states of which you are aware, your own mental states stand out by their phenomenality and by their apparent inner location. If one of those states only lacked the first of these salient features, it would seem to you not only that it fails to be yours but also (what could a thought be without a thinker?) that it is somebody else's.[7] Taking your phenomenology at face value, as one normally does when attributing ownership to a thought one is aware of, you would consequently endorse the delusional belief. If you finally tried to describe what it is like to experience such a good access to phenomenally unconscious thought vehicles, I take it that you would naturally describe those as images being flashed in your mind, or as voices inside you.

This explanation is accordingly charitable. The patient's beliefs can stand in relation to each other and to reality just as ours do, and be thus deemed "rational." The only trouble with the patient is that, because of some neurophysiological disorder, the piece of reality that lies inside him and determines his experiences is a bit strange. Some of the thoughts or thought-like processes which seem located in him are nevertheless phenomenally unconscious. Accordingly, these thought-like processes in him seem to convey someone else's thoughts. We do not need to posit any particular irrationality nor any odd belief-fixation mechanism to explain the

delusion. The patient notices a phenomenally unconscious thought-like process in him and understandingly attributes it to someone else.

Contrary to those we have reviewed so far, this hypothesis fits very well with the patients' reports. Indeed, it allows us to take many of the patient reports at face value, at least as far as their phenomenology is concerned. This is particularly striking with the patient quoted by Mellor (1970), who reported thoughts being flashed in his head like pictures on a screen, but reports of this kind are actually quite frequent. Patients often try to characterize their inserted thoughts in terms of some common vehicle for thoughts that is not considered phenomenally conscious but that somehow got to be inside their head. Such unconscious vehicles can be pictures on a screen like in the case above; they can also be mere information (''I wonder if that's me . . . it felt like a piece of information'' says a patient quoted by Hoerl, 2001, p. 190), waves (''it is like being stuck on the same wavelength as people'' says a patient quoted by Frith, 2005, 170), or, as we will come back to, voices (Frith's patient already quoted for example continues: ''I could hear everybody in my mind''[8]).

Many phenomenologically oriented psychiatrists consistently observe that schizophrenia seems to involve a problem of subjectivity and an alteration of the relationship between reflexivity and phenomenal consciousness. Most strikingly, Parnas and Sass (2001) notice that subjects with schizophrenia are often ''hyper-reflexive,'' and they relate this hyper-reflexivity to a deficit in the first-order aspects of consciousness that they characterize as a lack of ''presence'' and ''affectability.'' More specifically, they describe a young patient, Robert, who ''so to speak, witnessed his own sensory processes rather than living them. It applied to most of his experiences in that, instead of living them, he experienced his own experiences'' (Parnas & Sass, 2001, p. 105). Experiencing mental states or sensory processes in oneself that one does not live is exactly what I would call having a good reflexive awareness of states that lack (first-order) phenomenality. Parnas and Sass (2001) also notice that the experiences of subjects with schizophrenia manifest a strange form of spatialization:

> They describe their thoughts or feelings in physical terms, as if possessing an object-like spatial quality (''my thoughts are dense and encapsulated'') or locate them spatially (''my thoughts feel mainly in the right side of the brain'' or ''it feels as if my thoughts were slightly behind my skull''). One patient reported that her ''experiential point of perspective'' (presumably her experiential ''I pole'') felt ''as if'' spatially ''shifted some centimetres behind'' (she had a feeling as if she looked at the world somehow ''more from behind''). (Parnas & Sass, 2001, p.107–108)

The present hypothesis predicts that the subjectivity and (first-order) phenomenality of some of the patient's thoughts are replaced by a spatial analogue (apparent inner location), which nicely fits the reports of these patients.

As far as I can tell, this hypothesis is also consistent with our independent knowledge of schizophrenia, and it is no more speculative than the alternative hypotheses we considered. It would, however, benefit from an empirical confirmation, and in particular from a confirmation in controlled interviews aimed at determining the precise phenomenology of the patients. In the absence of such a study, I submit the following working hypothesis: we already have an important,

albeit implicit, database recording the patients' phenomenology; it consists in the description and in the clustering of the various neighboring symptoms, that is, of the symptoms that are characteristic of the same subtype of patients suffering from schizophrenia. It has been noticed in particular that thought broadcasting, alien voices and thought insertion seem characteristic of a subtype of schizophrenia (Minas et al., 1992). My account can readily explain this grouping.

Take alien voices first. My account suggests that inserted thoughts are similar to alien voices in many respects. In the same way, voices are vehicles of thought which are not phenomenally conscious but of which the subject is aware. Note, by the way, that the auditory phenomenology is not only characterized by a specific sensory quality. It differs from the phenomenology of other senses by the kind of sensori-motor contingencies it involves. Specifically, we seem to have a relatively robust and easy awareness of sounds but not of visual or tactile objects. Even when they do not seem to be in us, sounds seem to surround us. My hypothesis then predicts that alien voices are somehow continuous with inserted thoughts. The term 'voice' would often be used to describe some vehicles of thoughts of which the patient has a good awareness, irrespective of the sensory quality associated with such vehicles. Inserted thoughts and voices would differ in degree rather than in kind. Inserted thoughts would be inner voices, and alien voices would be outer inserted thoughts.

This prediction has received important confirmations. First, alien voices and other so-called auditory hallucinations in schizophrenia do not always have a genuine auditory phenomenology. Some congenitally deaf patients have, for example, reportedly complained of alien voices (Bergner, 2000; Feu & McKenna, 1999). Systematic studies of the phenomenology of auditory hallucinations have revealed that in many cases the patient does not consider this phenomenology to be auditory. Miller (1996) found that only 10 out of the 50 patients he questioned attributed an ''auditory phenomenology'' to their voices. In a much broader Internet study comparing the phenomenology of passive (intrusive) thoughts, active thoughts, and voices among OCD patients, schizophrenics, and healthy controls, Moritz and Larøi (2008) found out that while many healthy subjects and OCD patients assume that some auditory phenomenology is associated with their intrusive thoughts, many patients suffering from schizophrenia consider their ''voices'' to be unreal and not, or not very much, ''acoustic.'' Such a result is consonant with the fact that many neuroimaging studies have not found language production areas to be involved in auditory hallucinations and have observed activations in the auditory cortex only in a subgroup of patients (see Moritz & Larøi, 2008, for a review). Second, many studies have also shown that alien voices are not always heard outside: more than one third of the patients feel them to be ''inside the head'' and many (between 10% and 40% depending on studies) report them as being sometimes inside and sometimes outside the head. The location is independent of the severity of the symptoms (Copolov, Trauer, & Mackinnon, 2004; Junginger & Frame, 1985).

What about thought broadcasting? For healthy subjects, thoughts which seem inside their ego boundaries are also thoughts which seem to be theirs. This is not the case in schizophrenia. Patients suffering from schizophrenia find themselves with

thoughts in them which are not their thoughts. It is no wonder then that they might come to believe that, by parity of reasoning, their own thoughts might be broadcast in other people's minds as well. This would explain why schizophrenics seem to lose their sense of mental privacy, and it would nicely explain thought broadcasting.

5.1. Schizophrenia and the Two Dimensions of the Mind

I have criticized traditional accounts of thought insertion: they all assume that inserted thoughts are phenomenally conscious, accessible through introspection, and ultimately subjective, and they all end up explaining thought insertion in terms of agency. I have proposed an alternative explanation according to which thought insertion results from a disorder of ownership or subjectivity, itself stemming from the phenomenally unconscious character of a thought that is, in other respects, similar to our usual phenomenally conscious thoughts.

The problems encountered by traditional accounts of thought insertion and the availability of an alternative account should, I contend, suggest an important shift of emphasis in the cognitive study of schizophrenia. Basing his model on an analysis of delusions of control, Frith gave agency and action control a central role in the explanation of the symptoms of schizophrenia. It is time to explore deficits in subjectivity or ownership and phenomenal consciousness rather than agency.

The failure of the traditional accounts and the availability of an alternative to them also points toward a broader conclusion. It not only shows that the Cartesian principle is not threatened by the phenomenon of thought insertion; it shows that the principle is not threatened because there can be empirical dissociations between two sets of features which are very close and which are normally coinstantiated, but which are very different. The first set of feature comprises subjectivity, phenomenality, and reflected-on subjectivity and phenomenality. The second set of features comprises apparent inner location, reflected-on inner location, and the form of good reflexive awareness and availability to reflexive awareness that underlies phenomenal consciousness.

These two sets of features delineate two dimensions of the mind (Table 1). The first is essentially tied to phenomenality. It gathers properties that have something phenomenologically intrinsic. The second dimension, on the other hand, is

Table 1 Two dimensions of the mind. Even though they are usually associated, the phenomenal and the "spatial" or functional dimension of the mind can come apart and should not be confused.

The phenomenal mind	The spatial mind
phenomenal consciousness	immediate (?) availability to reflexive awareness
subjectivity (apparent ownership)	good availability to reflexive awareness (apparent inner location)
reflected-on phenomenal consciousness	immediate (?) reflexive awareness
reflected-on subjectivity	good reflexive awareness (reflected-on inner location)

apparently relational or functional in the broad sense. Along this second dimension, mentality and consciousness are essentially a matter of cognitive access and reflexivity, and ownership is essentially a matter of inner location. We can call these two dimensions the "phenomenal" and the "spatial" dimensions of the mind, respectively. Both dimensions are part of the common-sense view of the mind, and I take it that both dimensions inform more sophisticated theorizing about the mind.[9] It has been a common ambition of philosophers of mind to try to reduce the phenomenal dimension to the spatial dimension. The pervasive character of the classical interpretations of thought insertion and their common shortcomings seem to indicate that this ambition is not, however, without a risk. When we get into thinking about the mind, we seem all too prone to assimilate the phenomenal dimension to the spatial dimension, and to neglect the features that make of our mental states something we live subjectively, rather than something that merely "appears inside us." We are all too prone, in more metaphorical terms, to confuse our real self with what a patient quoted by Kimura (1997, pp. 331-3) calls a "'virtual self', a mere 'topological translocation' of the real self . . . that normally goes in parallel."

Notes

[1] "Which of those activities [doubting, understanding, willing, unwilling, imagining, feeling]," he asks in the second meditation, "is distinct from my thinking? Which of them can be said to be separate from myself? The fact that it is I who am doubting and understanding and willing is so evident that I see no way of making it any clearer" (Descartes, 1641/1985, p. 19)."

[2] Like most philosophers we will mention here, Coliva (2002) does not so much talk about the Cartesian principle, consciousness and subjectivity as about the immunity principle, "accessibility through introspection" or "first-person knowledge," and self-ascription based on introspection. I will neglect this difference as it is not relevant for what follows.

[3] See Pitt (2004) and Siewert (1998, section 8) for convincing, chapter-length, defenses of this claim.

[4] One might argue that the so-called "global workspace theory" supports such a link. However, defenders of this theory explicitly admit that it bears on forms of consciousness that are constitutively tied to reportability (Dehaene & Changeux, 2004; Dehaene, Changeux, Nacchache, Sackur, & Sergent, 2006; Kouider, Dehaene, Jobert, & Bihan, 2006), so it is not obvious at all that it bears on phenomenal consciousness. Other works on consciousness suggest that phenomenal consciousness does not require a good reflexive awareness (Block, 2005; Lamme, 2003, 2004; Zeki & Ffyche, 1998).

[5] Rosenthal used to claim that the entailment is a priori, but he now tends to claim that the link between reflexive awareness and consciousness or subjectivity is merely nomological and to present his higher order theory as an empirically based, rather than a conceptually based, theory.

[6] Some representationalists seem to claim that experiences are *wholly* diaphanous. I take it that they cannot do that while taking the subjectivity of experience seriously. I thank an anonymous referee for drawing my attention to this point.

[7] Interestingly, this kind of explanation of the step from "non-self attribution" to "other attribution" is not available to the agency theorist. For if it is plausible that a thought that is

not mine must be someone else's, it is not likewise plausible that a thought that is not caused by me must be caused by someone else rather than by no one.

[8] Cermolacce, Naudin, and Parnas also quote a patient, Thomas, who says "it is as if I was projected by voices which are not mine . . . I feel like he had left something inside my head" (2007, p. 708).

[9] This claim parallels Chalmers' (1996, p. 16) claim that mental terms have a double life, referring both to psychological and to phenomenal states or processes.

References

Bell, M., Fiszdon, J., Richardson, R., Lysaker, P., & Bryson, G. (2007). Are self-reports valid for schizophrenia patients with poor insight? Relationship of unawareness of illness to psychological self-report instruments. *Psychiatry Research*, *151*, 37–46.

Bergner, J. (2000). "Hearing voices" by a congenitally deaf schizophrenic patient. *Psychiatrische Praxis*, *27*, 49–50.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Science*, *6*, 237–242.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*, 227–287.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, *9*, 46–52.

Campbell, J. (1998). Le modèle de la schizophrénie de Christopher Frith [Christopher Frith's model of schizophrenia]. In H. Grivois & J. Proust (Eds.), *Subjectivité et conscience d'agir: Approches cognitive et clinique de la psychose* (pp. 99–113). Paris: PUF.

Campbell, J. (1999). Schizophrenia, the space of reasons and thinking as a motor process. *The Monist*, *82*, 609–625.

Campbell, J. (2004). The ownership of thoughts. *Philosophy, Psychology, Psychiatry*, *9*, 35–39.

Cermolacce, M., Naudin, J., & Parnas, J. (2007). The "minimal self" in psychopathology: Re-examining the self-disorders in the schizophrenia spectrum. *Consciousness and Cognition*, *16*, 703–714.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.

Coliva, A. (2002). Thought insertion and immunity to error though misidentification. *Philosophy, Psychiatry & Psychology*, *9*, 27–34.

Copolov, D., Trauer, T., & Mackinnon, A. (2004). On the non-significance of internal versus external auditory hallucinations. *Schizophrenia Research*, *69*, 1–6.

Dehaene, S., & Changeux, J.-P. (2004). Neural mechanisms for access to consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences III: Vol. 14* (pp. 1145–1159). Cambridge, MA: MIT Press.

Dehaene, S., Changeux, J.-P., Nacchache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *95*, 529–534.

Descartes, R. (1985). Meditations on first philosophy. In J. Cottingham & D. Murdoch (Eds.), *The philosophical writings of Descartes* (pp. 1–62). Cambridge: Cambridge University Press. (Original work published 1641).

Feu, M. D., & McKenna, P. (1999). Prelingually profoundly deaf schizophrenic patients who hear voices: A phenomenological analysis. *Acta Psychiatrica Scandinavica*, *99*, 453–459.

Freeston, M. H., & Ladouceur, R. (1999). Exposure and response prevention for obsessive thoughts. *Cognitive and Behavioral Practice*, *6*, 362–383.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove: Lawrence Erlbaum Associates.

Frith, C. D. (2005). The neural basis of hallucinations and delusions. *Comptes Rendus Biologies*, *328*, 165–175.

Frith, C. D., & Done, D. J. (1988). Towards a neuropsychology of schizophrenia. *British Journal of Psychiatry*, *153*, 437–443.

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: John Benjamins.

Graham, G., & Stephens, G. L. (2000). *When self consciousness breaks: Alien voices and inserted thoughts*. Oxford: Oxford University Press.

Henry, M. (1990). *Phénoménologie matérielle* [Material phenomenology]. Paris: Épiméthée PUF.

Hoerl, C. (2001). On thought insertion. *Philosophy, Psychiatry & Psychology*, *8*, 189–200.

Hoffman, R. E. (1986). Verbal hallucinations and language production processes in schizophrenia. *Behavioral and Brain Sciences*, *9*, 503–548.

Hume, D. (1978). *A treatise on human nature* (2nd ed.). New York: Oxford University Press. (Original work published 1740).

James, W. (1983). *The principles of psychology*. Cambridge, MA: Harvard University Press. (Original work published 1890).

Jaspers, K. (1997). *General psychopathology* (2nd ed., J. Hoenig & M. W. Hamilton, Trans.). Baltimore, MD: JHU Press. (Original work published 1913).

Junginger, J., & Frame, C. (1985). Self-report of the frequency and phenomenology of verbal hallucinations. *The Journal of Nervous and Mental Disease*, *173*, 149–155.

Kimura, B. (1997). Cogito et je [Cogito and I]. *L'évolution psychiatrique*, *6*, 335–348.

Kouider, S., Dehaene, S., Jobert, A., & Bihan, A. L. (2006). Cerebral bases of subliminal and supraliminal priming during reading. *Cerebral Cortex, 17*, 857–875.

Kriegel, U. (2003). Consciousness as intransitive self-consciousness: Two views and an argument. *Canadian Journal of Philosophy*, *33*, 103–132.

Kriegel, U. (2004). Consciousness and self-consciousness. *The Monist*, *87*, 185–209.

Kriegel, U. (2006). The same-order monitoring theory of consciousness. In U. Kriegel & K. W. Williford (Eds.), *Self-representational approaches to consciousness* (pp. 143–170). Cambridge, MA: MIT Press.

Lamme, V. A. F. (2003). Why attention and awareness are different. *Trends in Cognitive Sciences*, *7*, 12–18.

Lamme, V. A. F. (2004). Separate neural definitions of visual consciousness and visual attention: A case for phenomenal awareness. *Neural Networks*, *17*, 861–872.

Lichtenberg, G. C. (1990). *Aphorisms* (R. J. Hollingdale, Trans.). London: Penguin. (Original work published 1806).

Mellor, C. (1970). First rank symptoms of schizophrenia: 1. The frequnncy in schizophrenics on admission to hospital; 2. Differences between individual first rank symptoms. *British Journal of Psychiatry*, *117*, 15–23.

Miller, L. J. (1996). Qualitative changes in hallucinations. *The American Journal of Psychiatry*, *153*, 265–267.

Minas, I., Stuart, G., Klimidis, S., Facdson, H., Singh, B., & Copolov, D. (1992). Positive and negative symptoms in the psychoses: Multidimensional scaling of SAPS and SANS items. *Schizophrenia Research*, *8*, 143–156.

Moritz, S., & Larøi, F. (2008). Differences and similarities in the sensory and cognitive signatures of voice-hearing, intrusions and thoughts. *Schizophrenia Research*, *102*, 96–107.

Mullins, S., & Spence, S. A. (2003). Re-examining thought insertion semi-structured literature review and conceptual analysis. *The British Journal of Psychiatry*, *182*, 293–298.

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.

Parnas, J., & Sass, L. A. (2001). Self, solipsism, and schizophrenic delusion. *Philosophy, Psychiatry & Psychology*, *8*, 101–120.

Pitt, D. (2004). The phenomenology of cognition, or, what is it like to think that p? *Philosophy and Phenomenological Research*, *69*, 1–36.

Purdon, C. (2004). Empirical investigations of thought suppression in OCD. *Journal of Behavior Therapy and Experimental Psychiatry*, *35*, 121–136.

Purdon, C., Rowa, K., & Antony, M. M. (2005). Thought suppression and its effects on thought frequency, appraisal and mood state in individuals with obsessive-compulsive disorder. *Behaviour Research and Therapy*, *43*, 93–108.

Rachman, S., & da Silva, P. (1978). Abnormal and normal obsessions. *Behaviour Research and Therapy*, *16*, 233–248.

Rosenthal, D. M. (1997). A theory of consciousness. In N. Block, O. Flanagan, & G. Guzeldere (Eds.), *The nature of consciousness* (pp. 729–753). Cambridge, MA: MIT Press.

Salkovskis, P. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy*, *25*, 571–583.

Salkovskis, P. (1989). Cognitive-behavioural factors and the persistence of intrusive thoughts in obsessional problems. *Behaviour Research and Therapy*, *27*, 677–682.

Shoemaker, S. (1968). Self-reference and self-awareness. *Journal of Philosophy*, *65*, 555–567.

Siewert, C. (1998). *The significance of consciousness*. Princeton, NJ: Princeton university press.

Synofzika, M., Vosgeraub, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*, 219–239.

Zeki, S., & Ffyche, D. (1998). The Riddoch syndrome: Insights in the neurobiology of conscious vision. *Brain*, *121*, 24–45.