

Thought Insertion without Thought

Shivam Patel
Florida State University

Forthcoming in *Review of Philosophy and Psychology*
This is a preprint. Please cite the published version.

Abstract: There are a number of conflicting accounts of thought insertion, the delusion that the thoughts of another are inserted into one's own mind. These accounts share the common assumption of *realism*: that the subject of thought insertion has a thought corresponding to the description of her thought insertion episode. I challenge the assumption by arguing for an anti-realist treatment of first-person reports of thought insertion. I then offer an alternative account, *simulationism*, according to which sufferers merely simulate having a thought inserted into their heads. By rejecting realism, the paper undermines a widespread explanatory framework that unites otherwise competing cognitive models of thought insertion.

Word Count: 9,745

1. Introduction

Thought insertion is the delusion that the thoughts of another are inserted into one's own mind. The phenomenon is predominately found in people with schizophrenia, but is also present in those with mood disorders and acute psychosis (Peralta & Cuesta, 1999). The literature is marked by a number of conflicting accounts of thought insertion. Uniting these accounts is an assumption I will call *realism*, according to which the subject of thought insertion has a thought corresponding to the description of her thought insertion episode. Once it is granted that the subject has a thought, the explanatory task becomes determining how a thought had by the subject could seem to be someone else's thought. Existing accounts of thought insertion are various attempts at finding the processes that cause sufferers to take their own thought to be the thought of another.

In this paper, I reject the assumption of realism and thereby challenge the explanatory framework within which existing accounts of thought insertion have been developed. Section 2

illustrates how realism shows up in the literature. Section 3 makes precise the nature of the disagreement between realism and anti-realism. Section 4 argues that we have no reason to assume that realism is true. Section 5 makes the case that thought insertion episodes are analogous to dreams: just as dreaming that I am thinking does not involve actually thinking, so too having an episode in which a thought is inserted in my head does not involve actually having a thought. These arguments bring into view an alternative to realism, which I call *simulationism*, according to which sufferers of thought insertion do not have a thought, but, as with dreams, merely simulate having a thought inserted into their minds. After describing this alternative in Section 6, in Section 7 I illustrate how the view transforms the explanatory landscape of thought insertion. I diagnose the appeal of realism in Section 8 before concluding in Section 9.

2. Realism about Thought Insertion

Realism is at work in treatments of first-person reports of thought insertion. Authors often start by citing the following reports:

I look at 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).

Thoughts come into my head like ‘Kill god’. It’s just like my mind working, but it isn’t. They come from this chap, Chris. They’re his thoughts. (Frith, 1992, p. 66)

I didn’t hear these words as literal sounds, as though the houses were talking and I were hearing them; instead, the words just came into my head – they were ideas I was having. Yet I instinctively knew they were not *my* ideas. They belonged to the houses... (Saks, 2007, p. 27)

The assumption of realism is apparent in how authors explicate these reports. Consider Campbell (1999):

A patient who supposes that thoughts have been inserted into his mind by someone else is right about which thoughts they are, but wrong about whose thoughts they are.... The thought inserted into the subject’s mind is indeed in some sense his...He has, for example, some especially direct knowledge of it. On the other hand, there is, the patient insists, a sense in which the thought is not his, a sense in which the thought is someone else’s... (Campbell, 1999, pp. 609-610)

In claiming that subjects of thought insertion are “wrong about whose thoughts they are” but “right about which thoughts they are,” Campbell assumes that they have a thought corresponding to their description. Thus, according to Campbell, Frith’s subject actually has the thought *Kill god*. This realist way of reading first-person reports is present in a wide array of authors who otherwise disagree about thought insertion:

What patients seem to be implying...is that certain thoughts that appear in their own stream of consciousness are not experienced as their own... (Lopez-Silva, 2019, p. 326)

...the reports are extremely puzzling. How could anyone really think that thoughts woven into their stream of consciousness belonged to someone else? (Parrott, 2017, p. 40)

...in thought insertion, the subject is aware of the content of a thought in a first-personal way but does not acknowledge that the thought is her own...it is a puzzle for our philosophical conception of what it is to have a thought. (Bortolotti and Broome, 2009, p. 215)

These subjects seem to believe some things while claiming, at the same time, that the relevant belief is not theirs. This is puzzling...What we need is an account of the fact that subjects with thought insertion disown certain mental states that, nonetheless, they admit to have. (Fernandez, 2010, p. 68)

These authors all agree that thought insertion presents a “puzzle.” The puzzle is: how can a sufferer of thought insertion be possessed of a thought, but claim that it is the possession of another? The puzzle presupposes the truth of realism. Thus, Lopez-Silva and Parrott both imply that the subject has thoughts within their “stream of consciousness,” Bortolotti and Broome state that the subject is “aware of the content” of her thought, and Fernandez wonders how a subject can “have a mental state” but also “disown” it.

In addition to informing the analysis of first-person reports, realism also shapes explanations of thought insertion. Theorists tend to adopt an explanatory framework that appeals to two types of variables, a *thought variable* and a *process variable*. The thought variable is a thought that corresponds to the description of a subject’s thought insertion episode, e.g., the thought *Kill god*. The process variable is a process that operates on the thought variable and issues

in thought insertion episodes. Existing theories agree on the nature of the thought variable, but disagree about the process variable.

A popular debate concerns whether the process variable issues in a missing sense of being *the agent* of a thought or whether it issues in a missing sense of being *the owner* of a thought. According to *agency accounts*, subjects of thought insertion have a thought but are missing a sense of being its agent (Frith (1992), Campbell (1999), Sousa and Swiney (2013), Gallagher (2015), Zahavi and Kriegel (2015), Sollberger (2014), Duncan (2019)). Campbell (1999) presents the most influential agency account, which adapts Frith's explanation of passivity phenomena in schizophrenia. According to Frith (1992), action is controlled, guided, and corrected by a set of comparators that compare actual states with predicted or intended states. In normal action, motor commands drive limb movement. In addition to this feedforward process, there is also a lateral process whereby a copy of the motor commands is generated and fed into a 'forward model', which computes a prediction of the sensory consequences of successfully executing the motor command. According to Frith, the sense of agency in normal action derives from a match between actual and predicted sensory feedback. However, according to Frith, if no prediction of sensory feedback is computed, subjects experience a loss of agency over their actions. For example, consider how Frith accounts for the case of *made action*, where someone experiences their own actions as controlled by another. Motor commands drive a hand movement, say, but unlike in the case of normal action, no copy of the motor command is generated, and so there is no prediction of the upcoming hand movement. According to Frith, this failure to predict the upcoming hand movement results in the hand movement being felt as 'alien'.

Campbell (1999) applies this account to thought insertion. According to Campbell, background beliefs, desires, and so on, generate a command to think a thought, which drives the

thinking of the thought. According to Campbell, a “match at the monitor between the thought you have introspective knowledge of and the efferent copy [i.e., prediction] [is] responsible for the sense of being the agent of that thought” (p. 617). However, in the case of thought insertion, the prediction is not present, and so the subject does not have a sense of being the agent. On Campbell’s view, subjects of thought insertion have a sense of owning a thought, since they have introspective access to a thought in their minds, but they do not have a sense of being the agent of the thought due to a breakdown of the prediction. Campbell’s explanation thus appeals to a thought variable over which his preferred process variable – *breakdown of prediction* – generates thought insertion.

Even though Campbell’s model is more than thirty years old, it has continued to be influential, shaping models put forward by Gerrans (2001), Langland-Hassan (2008), and, more recently, predictive processing theorists (e.g., Sterzer et al. (2016)).¹ Although agency accounts remain popular, a number of authors argue that the problem is not with sufferers’ sense of agency (Bortolotti and Broome (2009), Howell and Thompson (2017), Mathieson (2023)). For example, Howell and Thompson (2017) argue that agency accounts fail to distinguish between thought insertion and uncontrollable, intrusive thoughts. At least some of these authors go on to argue that the problem in thought insertion lies instead in a lack of a sense of ownership over thoughts. For example, Bortolotti and Broome (2009) argue that since subjects of thought insertion fail to self-ascribe the inserted thought, they do not have a sense of ownership of the thought (for similar views, see Billon (2013) and Martin and Pacherie (2013)). Ownership accounts thus disagree with agency accounts about whether thought insertion arises from a breakdown in agency-generating

¹ See Graham and Stephens (2003) for an alternative agency account. Though Graham and Stephens agree with Campbell regarding the thought variable, they disagree regarding the process variable, instead explaining thought insertion in terms of a process of failing to make a thought cohere with one’s other intentional states (p. 102).

processes. For our purposes, however, what is important is that both accounts agree that thought insertion involves a thought variable.

Though the debate between agency and ownership theorists takes up much of the existing literature, some more recent authors have attempted to move past the standard dialectic. Both agency and ownership accounts view thought insertion as involving the *absence* of something normally present in thought – either a sense of agency or a sense of ownership. More recently, however, some have argued that something is *added* in the case of thought insertion. For example, Mathew Parrott (2017) argues that in episodes of thought insertion, someone has normal first-person awareness of a thought, but, in addition, has the sense that her state of awareness is not ordinary first-person awareness. A similar style of view is found in the phenomenological literature as well. Mathieson (2023) claims that thought insertion is not “a condition where some features of phenomenal consciousness are present while others are lacking [e.g., a sense of agency or ownership]....” Rather, according to Mathieson, thought insertion involves a disturbance in “for-me-ness,” the minimal form of self-awareness we have in thought (see also Henriksen, Parnas, & Zahavi 2019). These more recent accounts seek to characterize ways in which the sufferer of thought insertion undergoes a fundamental disturbance of thought awareness without appealing to the absence of some normal feature of such awareness. But even these accounts agree with both agency and ownership accounts that explanations of thought insertion involve appeal to a thought variable.

Finally, the assumption of realism also shows up in both analytic and phenomenological discussions of how thought insertion bears on self-consciousness. In the analytic tradition, *immunity to error with respect to the first-person* is the thesis that one cannot introspect a thought while also being in error about whose thought it is (Shoemaker, 1968). A number of authors have

taken thought insertion to be a counterexample to the immunity thesis, since in thought insertion one introspects a thought (e.g., the thought *Kill god*), but is in error about whose thought it is (e.g., by attributing the thought to Chris) (see, e.g., Campbell (1999) and Hu (2017)). In response, there have been attempts to show that thought insertion is not a counterexample to the immunity thesis (see, e.g., Coliva (2002) and Palmira (2020)). The phenomenological tradition contains a related discussion about whether thought insertion is a counterexample to the thesis that phenomenal consciousness involves a sense of mineness (Zahavi (2018)). Some argue that thought insertion is a counterexample by construing it as lacking a sense of mineness (Guillot (2017) and Howell and Thompson (2017)). In response, Zahavi and colleagues have attempted to save the thesis by claiming that thought insertion does involve a sense of mineness, but in a disrupted form (Henriksen, Parnas, & Zahavi 2019). For our purposes, what is important is that both sides of these disputes assume the truth of realism: that subjects of thought insertion have a thought concerning which they may or may not have some relevant form of self-consciousness or sense of mineness.

The assumption of realism therefore shows up in the literature in three ways: it is illustrated in analyses of first-person reports, it is presupposed in explanations of thought insertion, and it informs discussions of self-consciousness. In the next Section, I make precise the difference between realism and the kind of anti-realism I argue for in Sections 4 and 5.

3. Framing the Disagreement between Realism and Anti-Realism

Let's start with dreams. Dreams are such that if I dream that I am F-ing, then the dream itself does not involve actual F-ing. For example, if in my dream I run a mile, my dream does not

involve actual mile running. Now take hallucinations. Hallucinations are such that if I hallucinate an F, then the hallucination itself does not involve the actual presence of an F. For example, if in my hallucination many pink rats are crawling on me, the hallucination itself does not involve the presence of actual pink rats. Of course, it could be that I am actually running a mile *while* I am dreaming of running a mile or that pink rats are actually present *while* I hallucinate pink rats. But, the actual mile running and the actual pink rats would not be parts of the dream or the hallucination (Sosa, 2005).

I will argue that we should conceive of thought insertion episodes as akin to dreaming or hallucination. On this view, if I have a thought inserted in my mind by another in a thought insertion episode, then the episode itself does not involve my actually having a thought corresponding to the description of the episode. Just as dreaming that I am running a mile does not involve actual mile running so too having a thought insertion episode involving the thought *Kill god* does not involve the actual thought *Kill god*. Of course, it could be that the actual thought *Kill god* occurs *while* I am having the thought insertion episode just I could be actually running a mile *while* dreaming of running a mile. But the thought *Kill god* would not be part of the episode itself in just the way that actual mile running is not part of the dream of running a mile. Anti-realism claims that actual thoughts corresponding to first-person reports of thought insertion are not parts of thought insertion episodes. In contrast, realism claims that such thoughts are parts of thought insertion episodes.

We can make the disagreement between the realist and anti-realist more precise by distinguishing two aspects of a thought insertion episode. The first-person reports discussed above suggest that, from the point of view of the subject, there are two components of a thought insertion episode: what I will call a *thought component* and an *insertion component*. In characterizing the

thought component in the above reports, Mellor's subject speaks of thoughts, Saks refers to ideas, while Frith's subject provides a more specific characterization in terms of the thought's content. In contrast, the insertion component is the insertion event, which is variously described in the first-person reports. Saks says that the ideas do not belong to her, but "belong to the houses..." Mellor's subject says that Eammon Andrews "treats my mind like a screen and flashes thoughts onto it like you flash a picture." Notice that no theorist about thought insertion is a realist about the insertion component. All theorists agree that the insertion component is fictional: if I have a thought that is inserted by another in a thought insertion episode, then the episode does not involve an actual insertion event. The disagreement between the realist and the anti-realist concerns only the thought component.

One more clarification is in order. What do I mean by *thought*? There are more and less broad senses of "thought." In one sense, a thought might be any mental episode with propositional content. In this sense, imagining that one is flying through the sky would count as a thought, while itches or tickles would not. In another sense, however, thought might only encompass non-sensory states with propositional content. In this sense, the occurrent belief that it is raining would count as a thought, but imagining that one is flying through the sky would not. Fortunately, for our purposes, we do not need to determine in general terms what counts as a thought. Anti-realism claims only that thoughts *corresponding to first-person descriptions of thought insertion* are not parts of thought insertion episodes. For example, for Frith's subject the thought corresponding to the description is the thought *Kill god*, and so the anti-realist is only denying that this thought is a part of the episode. For simplicity's sake, in what follows I will often use "actual thought" as shorthand for "actual thought corresponding to the first-person description of the thought insertion episode."

In the next Section, I will show that there is little reason to favor realism. Since the assumption of realism has yet to be recognized, I will need to speculate about what might be said in its defense. Though this presents a significant limitation of the argument in Section 4, I will also provide a positive argument against realism in Section 5.

4. No Reason for Realism

I will now consider three possible defenses of realism. The first concerns properties of first-person reports, the second concerns properties of thoughts, and the third concerns empirical evidence from neurophysiology.

4.1. Commitment in First Person Reports

First-person reports of thought insertion seem to show commitment to actual thoughts. Such commitment not only seems to be present in the often-quoted examples above, but also in updated examples recently culled by Peter Langland-Hassan (forthcoming):

People from my past constantly interfere with my mind and insert things into it. It used to be manageable but it's becoming less so. Hoping Clozapine wipes out all of this nonsense, as it's driving me crazy. (from Joker, *Schizophrenia.com*)

I've been having this feeling that maybe this person that I know is inserting thoughts into my head or controlling them. Now I feel like my life isn't my own anymore and that I'm forever cursed to live with that person manipulating my mind. (from Fleur2576, *Schizophrenia.com*)

I have intrusive thoughts and inserted thoughts. Not sure what the difference is except inserted thoughts seem foreign or inserted and not from me, but are 'inserted' in my head from an external source like ET, chip, AI, God, quantum physics, parallel universes, past lives, etc. (from johnnyboy1, *Schizophrenia.com*)

These reports seem to express commitment to actual thoughts. For example, the reports from Joker and Fleur2576 both seem committed to thoughts – “things” and “thoughts,” respectively – while jonnyboy1 seems to be claiming that his actual thoughts seem to be “foreign.” Isn't this commitment enough to justify realism?

In Section 5 I will question whether sufferers are, in fact, expressing such a commitment. However, for the sake of argument, let us suppose that they are. The problem is that commitment alone cannot justify ascription of thoughts to subjects of thought insertion, since they also seem to commit themselves to entities and events – inserters and insertion events – that the realist would not dare posit. Joker saying that people from their past are inserting things in their mind does not give us reason to think that people are actually doing so. To the commitment condition the realist might add an extra, normalcy condition: we should posit actual thoughts but not inserters or insertion events because the former are normal while the latter not. The problem is that the commitment and normalcy conditions make the wrong prediction when it comes to dreams. Suppose I report the following from a dream:

(1) I had the thought that blueberries taste good.

By realist lights, the commitment and normalcy conditions would justify positing the actual thought that *blueberries taste good* as a part of my dream. After all, I seem to commit myself to the actual existence of the thought that *blueberries taste good* and thoughts are normal mental states. However, when it comes to dreams, the ascription of the actual thought is unjustified – the dream itself does not involve the actual thought that *blueberries taste good*. The commitment and normalcy conditions cannot guide the ascription of thoughts when it comes to dreams. Thus, if the realist posits an actual thought on the basis of the commitment and normalcy conditions, she is begging the question against the view that thought insertion episodes are akin to dreams. Of course, the commitment and normalcy conditions may provide good reason for the ascription of thought in *normal circumstances* (e.g., when you utter (1) while expressing your taste in fruit).

The problem is that the circumstances that surround first-person reports of thought insertion are far from normal.

4.2. Seeming Thought implies Thought

Another potential reason for realism stems from properties associated with thoughts. In particular, the realist might put forward the following conditional: if it seems to a subject S that S is having a thought, then S is *actually* having a thought. The principle might be motivated by appeal to analogies with pain and other phenomenal states. Thus, if it seems to S that S has a pain, then S *actually* has a pain, or if it seems to S that S experiences a reddish *quale*, then S *actually* experiences a reddish *quale*. Given that both the realist and anti-realist assume that it seems to the sufferer that she has an occurrent thought, it follows from the proposed conditional that the sufferer *actually* does have a thought, *contra* the anti-realist.

There are two problems with this argument. First, the conditional is commonly taken to apply to phenomenal states, e.g., pains, visual experiences, and the like (Hill, 2009), but not typically taken to apply to paradigmatic cognitive states such as occurrent thoughts. Second, it is plausible that there are cases in which it can seem that one is thinking a thought, but one is not *actually* thinking that thought. Consider a case of immersive imagination. Jill is writing a novel about a wife who murders her husband. For creative material, she imagines what it would be like to plot her own husband's murder. She becomes so absorbed in her imagining that she is taken aback and wonders whether she was *actually* plotting her husband's murder. It seems to Jill as if she is actually thinking about killing her husband – this is why she is left worried. But no matter how immersive her imagining, she is still only *imagining* thinking through her husband's murder,

not actually doing so. Unlike with phenomenal states, then, the principle seems false when applied to actual occurrent thought.

An objector might reject this line of reasoning if they hold a constitutive view of self-knowledge. According to constitutive views, given certain conditions *C*, *S* believes/thinks/etc. that *p* if and only if *S* believes (or judges) that she believes/thinks/etc. *p* (Coliva, 2016). Since the subject of thought insertion believes *that she thinks Kill god*, by the constitutive thesis the subject also thinks *Kill god*. Constitutivism thus seems to provide a framework that favors realism. However, first-person reports do not indicate that the sufferer believes *that she thinks Kill god*. Instead, all indications point to the sufferer having a belief with a far more complex content, namely, that *she thinks the thought Kill god is being inserted into her head by Chris*. However, plugging this belief state into the constitutive thesis would not entail that the subject thinks *Kill god*. Instead, it would entail only that the sufferer thinks that *the thought Kill god is being inserted into her head by Chris*. Moreover, the constitutive thesis is taken to apply when a specific set of conditions hold, one of which is that the subject is “lucid” (Coliva, 2016, p. 64). However, this lucidity requirement makes the constitutive thesis inapplicable to people with thought insertion.

4.3. Neurophysiological Evidence

Finally, the realist might argue that neuroimaging provides promising evidence in favor of realism. There is very little neuroimaging work on thought insertion. However, in a rare exception, Walsh et al. (2015) use fMRI to show a decrease in the activation of language areas of the brain during thought insertion-like events. Let’s suppose that the activation patterns observed by Walsh et al. are indicative of the presence of an actual thought. If this is so, then Walsh et al.

seem to provide some evidence in favor of the idea that thought insertion episodes involve actual thoughts.

The problem with using this study to support realism is that Walsh et al. did not directly study thought insertion episodes. They recruited highly suggestible participants and had them complete sentence stems (e.g., ‘The dog...’). In the experimental condition, participants were told that they would experience the fMRI engineer inserting a sentence completion into their minds (e.g., ‘...has brown fur’). This was supposed to mimic having a genuine thought insertion episode. The problem is that successful execution of the task seems to require that the subject *simulate* having a thought inserted into one’s mind. But *simulating* the fMRI engineer putting the thought ‘the dog has brown fur’ into my mind does not entail that I thereby *actually* have the thought *that the dog has brown fur*. It entails only that I simulate having the thought as part of a larger simulation of it being inserted into my mind. Moreover, this suggests that even if *genuine* cases of thought insertion implicate decreased activity in language areas of the brain, as Walsh et al. show for mock thought insertion, this would not distinguish between the subject actually having a thought versus merely simulating having a thought. Neuroimaging would not cut ice when it comes to vindicating realism.

5. A Reason to Reject Realism

My positive argument against realism will focus on properties of first-person reports. In reading *The Adventures of Sherlock Holmes*, we find the following sentence:

(2) I rose to go, but Holmes caught me by the wrist and pushed me back into my chair.

In a report of a dream, we might hear the following:

(3) I was in the state of Washington, and I saw a guy walking down the street with a sweatshirt on. (from Auld, et al., 1968, p. 421)

(2) and (3) are true despite the fact that there is no such person as Sherlock Holmes and the speaker was never in Washington.² The phenomenon is not relegated to tellings of fictions and dreams, but is also present in reports of symptoms that afflict those with schizophrenia. Consider the following report of hallucination:

I was on the street and it was a guy calling my name, I thought it were some random dudes on [*sic*] a coffee shop and ran away from them scared shitless. (from anon72351231, *Schizophrenia.com*³)

anon72351231's report is true even though there was no actual voice calling his name. The sentences that compose the report are true despite the fact that the object or properties mentioned in the sentences do not exist.

I suggest that the same holds for reports of thought insertion episodes. In Section 3, I distinguished between the insertion component and thought component of thought insertion episodes. At the level of the report, we can distinguish between the *sentences that express the insertion component* (I-sentences) and the *sentences that express the thought component* (T-sentences). I think it is clear that I-sentences are true despite the non-existence of insertion events. Consider a fragment from Mellor's subject:

(4) The thoughts of Eamonn Andrews come into my mind.

² There remains the leftover question about what accounts for the truth of (2) and (3). A common view is that (2) is not *actually* true, but true *in the fiction* (Woodward, 2011). The account I provide of thought insertion is compatible with an analogous position, according to which sentences describing thought insertion episodes are not *actually* true, but true *in the simulation* (see Section 6).

³ Retrieved on August 20, 2022 from <https://forum.schizophrenia.com/t/people-with-auditory-hallucinations-question/53126/19>

(4) remains true even if we suppose that there is no such person as Eamonn Andrews. Or consider a regimented fragment from johnnyboy1's report:

(5) Thoughts are inserted in my head from ET.

I suggest that (5) is true despite the fact that ET does not exist.

How do things stand with T-sentences? In the context of describing his thought insertion episode, Mellor's subject starts out with the following:

(6) I look at the window and I think that the garden looks nice and the grass looks cool.

Suppose that Mellor's subject did not actually have the thought that *the garden looks nice* and *the grass looks cool*. Instead, suppose that it was merely as if he was having these thoughts. There is no Eamonn Andrews, no insertion event, no thoughts, but it is merely as if there is an Eamonn Andrews, it is merely as if there is an insertion event, and it is merely as if there is a thought. In this situation, I suggest that (6) remains true in just the same way that the hallucination report remains true even if there are no voices. If this is so, however, then T-sentences are true despite the non-existence of actual thoughts. But this means it can be true that sufferers are undergoing thought insertion episodes even though they do not have actual thoughts. Therefore, *contra* realism, thoughts are not parts of thought insertion episodes.

The realist will resist this analysis of T-sentences. They will claim that if it is merely as if there is a thought, then sentences like (6) are *false*. This response requires a picture on which the sufferer of thought insertion is making a report with *objective purport*, a report about how things objectively are with herself, and not a report with merely *subjective purport*, a report about how things merely seem. A report with objective purport is false while a report with subjective purport is true if the subject reports that she has a thought, but it is merely as if she has a thought. There are several reasons to think that T-sentences do not have objective purport. Consider again a

situation in which it is merely as if there is an inserter, it is merely as if there is an insertion event, and it is merely as if there is a thought. There are two possible positions the realist can adopt if T-sentences have objective purport in this situation. The first option is to claim that both I-sentences and T-sentences are *false*, and the second option is to claim that while I-sentences are *true*, T-sentences are *false*.

According to the first option, thought insertion reports as a whole have objective purport. But a comparison with dream discourse suggests that reports of thought insertion do not have objective purport. There are significant similarities between dream and thought insertion reports. The awoken dreamer shows insight into the dream-status of her dream. The awoken dreamer does not believe that she actually flew through the sky, and so on. Nevertheless, the awoken dreamer's report uses language that is typically associated with reports that have objective purport. Consider the language of the following dream report:

I was in the state of Washington, and I saw a guy walking down the street with a sweatshirt on. It said "Booth Newspapers." The next thing I knew there were girls running in a track race. Then I was in a photography class here at Wayne and the instructor was looking at the pictures and analyzing them. (Auld, et al., 1968, p. 421).

Given that awoken dreamers show insight into the unreality of the dream, the most charitable interpretation is that their reports have *merely apparent* objective purport: though the reports use language associated with reports that have objective purport, they do not, in fact, have objective purport. The truth of dream reports does not turn on how things stand in the objective order.

I suggest that a similar line of reasoning applies to thought insertion reports. At least some sufferers of thought insertion show insight into the unreality of their insertion episodes. Thus consider Elyn Saks' (Section 2) and johnnyboy1's (Section 4.1) reports. johnnyboy1 shows insight in his use of scare quotes around the word "inserted," suggesting that he believes insertion events do not actually exist. Moreover, johnnyboy1 comments only on how things "seem[ed]" to him.

The most charitable interpretation is therefore that johnnyboy1's report does not have objective purport. Saks, herself a researcher at the intersection of law and schizophrenia, shows insight that there are neither insertion events nor inserters. Despite this insight into the unreality of her thought insertion episodes, Saks uses objective language just as the awoken dreamer does. For example, Saks says that she "instinctively knew they were not *my* ideas" but that they "[t]hey belonged to the houses...." The most charitable interpretation is that her report has *merely apparent* objective purport. Similar analyses are plausible in general. If this argument is on the right track, the realist cannot assume that both I-sentences and T-sentences have objective purport, and so cannot take up the first option.

One might object that I have relied on only two instances where subjects of thought insertion have insight into the unreality of their episodes. According to the objection, most sufferers of thought insertion do not have insight, and so most reports have objective purport. In other words, according to the objection, sufferers tend to *believe* that someone is inserting thoughts into their heads. In response, we need to distinguish between reports of thought insertion episodes, on the one hand, and reports of what is believed or not believed on the basis of thought insertion episodes, on the other. In the case of hallucination reports, for example, anon7235123 reports their hallucination by saying that they are having an experience of someone calling their name irrespective of whether they believe that there is someone calling their name or not. The hallucination report is a report of the *underlying experience* and does not reflect one's *belief* concerning the experience.

In the same way, I suggest that reports of thought insertion episodes are reports of the *underlying experience* of having a thought inserted into one's head, and not reports of *belief* concerning the experience. Consider Fleur2576, who reports "...this *feeling* that maybe this

person that I know is inserting thoughts into my head...,” suggesting that the report concerns her *feeling* or *experience* in undergoing thought insertion. Mellor’s subject also seems to be describing his underlying experience in analogizing thought insertion to the workings of a projector. These reports of thought insertion seem to be reports of experience, not belief. Moreover, the idea that thought insertion reports are reports of underlying experience is required to make sense of reports of those with insight. It would make sense for those with insight to use objective language despite their disbelief in inserters and insertion events only if the language were tracking their underlying experience and not their belief state. If first-person reports of thought insertion track experience and not belief, as I have been arguing, then the supposed fact that most sufferers believe that there are actually thought inserters is irrelevant to the purport possessed by their reports.

We have so far argued that the realist cannot hold that both I-sentences and T-sentences are false in the situation in which it is merely as if there is an inserter, insertion event, and thought. But perhaps the realist will take up the second option by allowing that in such situations I-sentences do not have objective purport and so are true, while nevertheless maintaining that T-sentences have objective purport and so are false. For example, according to this option, in such situations it is true that someone is inserting thoughts into johnnyboy1’s head, but false that he has a thought. The problem is that this view implies that the sufferer is using two different voices in giving one and the same report: in expressing the T-sentence, she is speaking with objective purport, while in expressing the I-sentence she is speaking with merely apparent objective purport. But it is deeply implausible that this is what goes on when a sufferer reports on their thought insertion episode. Suppose a psychiatrist is interviewing a patient and asks, “What was happening before your mother committed you?” The patient responds:

(7) (a) I had thoughts that the grass looked cool. (b) But Eamonn Andrews way over in England was inserting the thought into my head.

In the situation in which it is merely as if there is an inserter, insertion event, and thought, it would be strange for the psychiatrist to treat (7b) as true, but (7a) as false. If the sentence expressing the insertion component is deemed true, then so too should the sentence expressing the thought component. The realist should not adopt the second option, on which I-sentences are true but T-sentences are false, in a situation in which it is merely as if there is an inserter, insertion event, and thought.

I conclude that the realist cannot adopt an analysis of T-sentences according to which they are false in the situation in which it is merely as if there is a thought. This suggests that we should adopt a view on which T-sentences are true in such a situation. But this means that thoughts are not parts of thought insertion episodes: one can undergo a thought insertion episode without having an actual thought.

The realist may question the above argument on methodological grounds. In particular, they may doubt the move from facts about the *language* in which thought insertion episodes are reported to facts about *thought insertion episodes* themselves. According to this objection, the fact that a report has merely apparent objective purport has no bearing on the nature of the underlying state that is thereby reported. However, this seems wrong in the case of dreams. Consider the following dream report:

(8) I flew through the sky.

The truth of (8) does not turn on the layout of the objective order – whether I actually flew through the sky or not. But this makes sense only if dreaming that I flew through the sky is silent on whether one actually flew through the sky or not. If dreaming that I flew through the sky required that I actually flew through the sky, we should expect the truth of (8) to turn on how things stand in the objective order. But this just means that we can use properties about dream reports – that

(8) is true whether or not I actually flew through the sky – to infer properties about dreams – that dreaming that I flew through the sky is itself neutral on whether I actually flew through the sky. Moreover, the realist is in no position to claim that the features of a report have no bearing on the state reported. This is because the primary motivation for realism turns on features of first-person reports, namely, that they show commitment to thoughts (Section 4.1). The realist’s methodological objection falls flat.⁴

6. Simulationism

The argument presented in the previous Section suggests that thought insertion episodes are simulations: just as dreaming is the *simulation* of living through a scenario, and hallucination the *simulation* of hearing a voice, so too in cases of thought insertion one *simulates* having a thought of another person inserted into one’s mind. According to simulationism, thought insertion does not involve actually having a thought, but simulating having a thought inserted by another.⁵ A crucial part of simulationism is that the simulation takes ‘wide scope’. It is not the case that one actually has a thought that one simulates being inserted in one’s head. This would just be an iteration of realism. Nor is it the case that a simulated thought is actually being inserted by another. There are no insertion events. Nor again is it that one experiences a simulated thought as inserted

⁴ This Section is consonant with two claims of the phenomenological approach. Central to the phenomenological approach is the idea that first-person reports of thought insertion should, in general, not be given a literal interpretation (Henriksen, Parnas, and Zahavi, 2019). In line with this claim, this Section implies that thought insertion reports that are apparently about the objective world, such as Joker’s report, should not be interpreted literally as being about the objective world, but instead as referring only to how things are experienced. The Section also aligns with another claim made by the phenomenological approach: that people with schizophrenia often engage in ‘double-bookkeeping’, whereby subjects are able to distinguish between the objective world and the world of their own delusion (Parnas, Urfer-Parnas, and Stephensen, 2021). In arguing that reports of thought insertion refer to experience and not the objective world, I am suggesting that sufferers of thought insertion understand the difference between the objective and delusional world.

⁵ There has been much empirical work supporting connections between simulation and delusion in general (see, e.g., Gerrans, 2014).

by another. Thought insertion concerns thoughts, not *simulated* thoughts. Rather, to accept that the simulation takes ‘wide scope’ is to just accept that in thought insertion one simulates having a thought inserted by another.

What are simulations? The minimal condition on being a simulative state is that it is possible to be in a simulative state without actually being in the state it simulates. In being in the simulative state of having the thought *Kill god* inserted into one’s head, one is in a state that simulates having the thought inserted into one’s head. One is not in a state that involves actually having the thought inserted into one’s head. There remain a number of questions that need answering if one seeks a more complete account of simulative states. First, it is an open question whether simulative states of thought insertion are imaginings, mental images, or otherwise *sui generis*. There also remains a question concerning the functional role of simulative states. It is commonly noted that delusions give rise to inconsistent behaviors (Bortolotti, 2010). For example, a person who suffers from Capgras syndrome may claim that his wife has been replaced by a look-alike imposter and yet show tender affection to the person. Such inconsistent behavior may also be present in thought insertion, in which case the simulative state will need to possess a similarly mixed functional role. Finally, a number of sufferers of thought insertion report misery and distress (see, for example, the reports by Fleur2576 and johnnyboy1). It is an open question what exactly accounts for this distress. But one option is that thought insertion episodes are akin to living nightmares: intrusive, dream-like simulations of thoughts of others being inserted into one’s mind.⁶

⁶ The view offered here bears similarities to Gregory Currie’s view of thought insertion (Currie (2000)). On one reading of Currie’s view, in thought insertion a subject is in fact imagining another’s thought in an attempt at mindreading. But, according to Currie, since she loses agency with respect to the imagining, she experiences imagining another’s thought as actually having another’s thought. This version of Currie’s view is anti-realist, since it accounts for thought insertion without appeal to actual thoughts, but imagined thoughts. However, unlike the simulationist, Currie claims that imagining another’s thought plays an *explanatory role* with respect to thought insertion episodes: imagining another’s thought is the basis upon which one comes to treat the imagined thought as an

7. Explanatory Implications: Realism vs. Simulationism

The reader may question the significance of arguing that thought insertion episodes do not involve actual thoughts but rather simulations of thoughts. From this perspective, the distinction between actual thoughts and simulations of thoughts is too fine and cuts no theoretical ice. What difference could it make whether a subject of thought insertion actually has a thought or merely simulates a thought? In fact, quite a big difference. To understand the implications of rejecting realism and adopting simulationism, we need to distinguish between the *nature* or *constitution* of thought insertion episodes, on the one hand, and the *explanation* of those episodes, on the other.

The paper has so far argued that thought insertion episodes are not constituted by actual thoughts. In contrast, existing views of thought insertion assume that thought insertion episodes are so constituted (Section 2). For example, agency accounts assume that in thought insertion one loses a sense of agency over an actual thought, while ownership accounts claim that one loses a sense of ownership of an actual thought. However, if thought insertion episodes are not constituted by actual thoughts, as argued in Section 5, then agency and ownership accounts fail for the same reason: there are no actual thoughts concerning which one loses agency/ownership. Existing views are mistaken about the nature of thought insertion episodes.

There is also a major difference between how the realist and simulationist will explain thought insertion episodes. Both realists and simulationists will appeal to molecular (e.g., overactive dopamine processes), neural (e.g., default-mode network dysregulation), psychosocial (e.g., childhood abuse or stress), and at least some cognitive factors (e.g., abnormal salience processing, activation of the default mode network). However, there is one factor that the

actual thought. In contrast, according to simulationism, thought insertions episodes just *are* simulations of having an inserted thought, and so simulations do not play an explanatory role with respect to such episodes.

simulationist will generally not cite in explaining thought insertion: actual thoughts corresponding to first-person descriptions of thought insertion episodes. In contrast, the realist must appeal to such thoughts, since, according to realist explanations, a thought undergoes an abnormal process that causes the subject to experience the thought as inserted (Section 2).

Of course, the simulationist *could* grant that the actual thought *Kill god* explains the ‘*Kill-god*’-episode. But the sense in which the thought would explain the episode for the simulationist differs significantly from the realist. For the simulationist the ‘*Kill god*’-episode might not have been caused by the thought *Kill god*, but by the thought *this is stressful*. For the simulationist, whether there is a match or mismatch between the content of the thought – *Kill god* or *this is stressful* – and the content of the episode – *Kill god* – has no explanatory significance whatsoever. Once it is granted that actual thoughts do not constitute thought insertion episodes, there is no longer motivation for requiring that the thought that explains a thought insertion episode match the content of the thought insertion episode. In contrast, from the perspective of realism, to suppose that the ‘*Kill god*’-episode occurs in the absence of the actual thought *Kill god* is incoherent.

Moreover, because there is no essential link between the content of the thought and the content of the thought insertion episode, it would be sheer coincidence that out of all possible thoughts the thought *Kill god* caused the ‘*Kill god*’-episode. From the perspective of simulationism, the thought *Kill god* causing the ‘*Kill god*’-episode would be coincidental in the same sense in which it would be coincidental that a hallucination of pink rats crawling on a wall was caused by actual pink rats. Just as we should not in general explain hallucinations of pink rats in terms of the presence of actual pink rats, so too we should not in general appeal to content-matching thoughts in accounting for thought insertion episodes. Overall, then, by rejecting realism

and adopting simulationism we undermine widespread assumptions about the nature and explanation of thought insertion: actual thoughts are not parts of thought insertion episodes and thought insertion episodes are not to be explained in terms of content-matching thoughts.⁷

8. Diagnosing Realism

The paper raises a broader question: why do theorists so often seek to understand thought insertion in terms of actual thoughts? The main attraction is that realist explanations conform to our *phenomenological* or *empathetic understanding* of thought insertion. Phenomenological understanding involves grasping what it is like to have an experience; it is something we have when we see someone stub their toe but lack when we think about bat sonar. As we have seen, realists appeal to thought and process variables in explaining thought insertion (Section 2). For example, Campbell claims that thought insertion episodes result from a failure to predict (process variable) a thought (thought variable). I suggest that Campbell's explanation provides scaffolding for a phenomenological understanding of thought insertion. In fact, Kendler and Campbell (2014) argue that explanations like Campbell's offer phenomenological understanding of otherwise "un-understandable" mental illnesses. Consider their discussion of Frith's explanation of made action from Section 2:

What is happening during these experiences? To anthropomorphize, in the first instance, when your arm is moving, a control monitor yells out, 'Is this one of ours?' Someone looks it up in the efferent copy bin and sure enough there is a copy of the 'lift your arm' command and the message goes back, 'One of ours!' So you have the subjective sense of it being your movement. In the second instance, no copy of the command is found in the efferent copy bin so the response comes back, 'Not one of ours!' You sense someone else moving your arm.... Assume we are in the future when this model has been confirmed by advanced neuroscience methods. You are interviewing Ms. Y who...reports extensive made-actions...Will you confront an un-understandable symptom beyond your empathic abilities?

⁷ I have focused on explanatory implications of adopting anti-realism. Anti-realism also has significant implications regarding the discussions of self-consciousness mentioned in Section 2. For example, if we adopt anti-realism, then the phenomenon of thought insertion has no bearing on the immunity thesis for the simple reason that thought insertion does not involve thoughts.

Kendler and Campbell answer ‘No’. According to Kendler and Campbell, states and processes implicated in Frith’s causal explanation – unexpectedness, other-attribution, arm movement – can be leveraged in simulations – a friend moving one’s arm unexpectedly – to gain an empathetic or phenomenological understanding of made action. Similarly, the states and processes that Campbell appeals to – unexpectedness, other-attribution, and actual thought – can be leveraged in simulations – a simulation of having someone else’s thought – to gain a phenomenological understanding of thought insertion. What gives the explanations intuitive force is just that they conform to our phenomenological understanding of the target.

In contrast, the explanatory resources provided by anti-realism do not offer phenomenological understanding of thought insertion. The anti-realist may claim, for example, that childhood abuse and default-mode network dysregulation converge to explain thought insertion. Of course, this explanation does provide *causal understanding* of thought insertion: it occurs because of the childhood abuse and default-mode network dysregulation. But the explanation does not provide phenomenological understanding, since understanding default-mode network dysregulation and childhood abuse will not give us a clue as to what it is like to experience a thought insertion episode. Unlike the realist, the anti-realist does not appeal to actual thoughts, and so does not explain thought insertion in a way that would scaffold phenomenological understanding.

But this is not a strike against anti-realism, since we should not insist that explanations of thought insertion provide phenomenological understanding. There are a number of mental illnesses whose explanation would not provide phenomenological understanding of the target. Consider manic episodes. Suppose we come upon a complete characterization of the causes of a manic episode for some particular person, which might include genetic predisposition, divorce,

stress, and lack of sleep. Someone without bipolar disorder might have a complete understanding of these underlying causes but will not thereby have a phenomenological understanding of what it is like to have a manic episode. My suggestion is that the same form of explanatory understanding we seek in the case of manic episodes should also be sought when it comes to thought insertion.

9. Conclusion

Explanations of thought insertion episodes are framed as solutions to a philosophically compelling puzzle: how could it be that a subject who has a thought experiences the thought as the thought of another person? What I have argued here is that there is no such puzzle, since thought insertion does not involve thought. This does not mean that there are no puzzles about thought insertion. It is genuinely puzzling how simulations of inserted thoughts arise, but it is puzzling in just the way it is puzzling how manic episodes, anxious fits, or the binge-purge cycle arise. There is no philosophically deep puzzle about thought insertion.

Declarations

Competing Interests

The author has no conflicts of interest.

References

- Auld, F., Goldenberg, G. M., & Weiss, J. V. (1968). Measurement of primary-process thinking in dream reports. *Journal of Personality and Social Psychology*, 8(4), 418–426.
- Billon, A. (2013). Does Consciousness Entail Subjectivity? The Puzzle of Thought Insertion. *Philosophical Psychology*, 26(2), 291–314.
- Bortolotti, L. (2010). *Delusions and Other Irrational Beliefs* (1st edition). Oxford University Press.
- Bortolotti, L., & Broome, M. (2009). A role for ownership and authorship in the analysis of thought insertion. *Phenomenology and the Cognitive Sciences*, 8(2), 205–224.
- Campbell, J. (1999). Schizophrenia, the Space of Reasons, and Thinking as a Motor Process. *The Monist*, 82(4), 609–625.
- Coliva, A. (2002). Thought Insertion and Immunity to Error Through Misidentification. *Philosophy, Psychiatry, & Psychology*, 9(1), 27–34.
- Coliva, A. (2016). *The Varieties of Self-Knowledge*. London: Palgrave.
- Currie, G. (2000). Imagination, Delusion and Hallucinations. *Mind & Language*, 15(1), 168–183.
- Doyle, S. A. C. (2009). *The Adventures of Sherlock Holmes* (Green ed. edition). Dover Publications.
- Duncan, M. (2019). The Self Shows Up in Experience. *Review of Philosophy and Psychology*, 10(2), 299–318.
- Fernández, J. (2010). Thought Insertion and Self-Knowledge. *Mind & Language*, 25(1), 66–88.
- Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia* (pp. xiv, 169). Lawrence Erlbaum Associates, Inc.
- Gallagher, S. (2015). Relations Between Agency and Ownership in the Case of Schizophrenic Thought Insertion and Delusions of Control. *Review of Philosophy and Psychology*, 6(4), 865–879.
- Gerrans, P. (2001). Authorship and Ownership of Thoughts. *Philosophy, Psychiatry, & Psychology*, 8(2), 231–237.
- Gerrans, P. (2014). *The Measure of Madness: Philosophy of Mind, Cognitive Neuroscience, and Delusional Thought*. Bradford Books.
- Graham, G., & Stephens, G. L. (Eds.). (2003). *Philosophical Psychopathology*. MIT Press.
- Guillot, M. (2017). I Me Mine: On a Confusion Concerning the Subjective Character of Experience. *Review of Philosophy and Psychology*, 8(1), 23–53.
- Henriksen, M. G., Parnas, J., & Zahavi, D. (2019). Thought insertion and disturbed for-me-ness (minimal selfhood) in schizophrenia. *Consciousness and Cognition*, 74, 102770.
- Howell, R. J., & Thompson, B. (2017). Phenomenally Mine: In Search of the Subjective Character of Consciousness. *Review of Philosophy and Psychology*, 8(1), 103–127.
- Hu, I. (2017). The Epistemology of Immunity to Error through Misidentification. *The Journal of Philosophy*, 114(3), 113–133.
- Kendler, K. S., and J. Campbell. 2014. “Expanding the Domain of the Understandable in Psychiatric Illness: An Updating of the Jaspersian Framework of Explanation and Understanding.” *Psychological Medicine* 44 (1): 1–7.
- Langland-Hassan, P. (2008). Fractured Phenomenologies: Thought Insertion, Inner Speech, and the Puzzle of Extranity. *Mind & Language*, 23(4), 369–401.
- Langland-Hassan, P. (forthcoming). Thought Insertion as Persecutory Delusion.

- López-Silva, P. (2019). Me and I Are Not Friends, Just Acquaintances: On Thought Insertion and Self-Awareness. *Review of Philosophy and Psychology*, 10(2), 319–335.
- Martin, J.-R., & Pacherie, E. (2013). Out of nowhere: Thought insertion, ownership and context integration. *Consciousness and Cognition: An International Journal*, 22, 111–122.
- Mathieson, D. (2023). Schizophrenic Thought Insertion and Self-Experience. *Review of Philosophy and Psychology*.
- Mellor, C. S. (1970). First rank symptoms of schizophrenia. *The British Journal of Psychiatry: The Journal of Mental Science*, 117(536), 15–23.
- Palmira, M. (2020). Correction to: Immunity, thought insertion, and the first-person concept. *Philosophical Studies*, 177(12), 3861–3861.
- Parnas, J., Urfer-Parnas, A., & Stephensen, H. (2021). Double bookkeeping and schizophrenia spectrum: Divided unified phenomenal consciousness. *European Archives of Psychiatry and Clinical Neuroscience*, 271(8), 1513–1523.
- Parrott, M. (2017). Subjective Misidentification and Thought Insertion. *Mind and Language*, 32(1), 39–64.
- Peralta, V., & Cuesta, M. J. (1999). Dimensional structure of psychotic symptoms: An item-level analysis of SAPS and SANS symptoms in psychotic disorders. *Schizophrenia Research*, 38(1), 13–26.
- Saks, E. R. (2008). *The Center Cannot Hold: My Journey Through Madness* (Reprint edition). Hachette Books.
- Shoemaker, S. S. (1968). Self-Reference and Self-Awareness. *Journal of Philosophy*, 65(October), 555–567.
- Sollberger, M. (2014). Making Sense of an Endorsement Model of Thought-Insertion. *Mind & Language*, 29(5), 590–612.
- Sosa, E. (2005). Dreams and Philosophy. *Proceedings and Addresses of the American Philosophical Association*, 79(2), 7–18.
- Sousa, P., & Swiney, L. (2013). Thought Insertion: Abnormal Sense of Thought Agency or Thought Endorsement? *Phenomenology and the Cognitive Sciences*, 12(4), 637–654.
- Sterzer, P., Mishara, A. L., Voss, M., & Heinz, A. (2016). Thought Insertion as a Self-Disturbance: An Integration of Predictive Coding and Phenomenological Approaches. *Frontiers in Human Neuroscience*, 10, 502.
- Walsh, E., Oakley, D. A., Halligan, P. W., Mehta, M. A., & Deeley, Q. (2015). The functional anatomy and connectivity of thought insertion and alien control of movement. *Cortex*, 64, 380–393.
- Woodward, R. (2011). Truth in Fiction. *Philosophy Compass*, 6(3), 158–167.
- Zahavi, D. (2018). Consciousness, Self-Consciousness, Selfhood: A Reply to some Critics. *Review of Philosophy and Psychology*, 9(3), 703–718.
- Zahavi, D., & Kriegel, U. (2015). For-me-ness: What it is and what it is not. In *Philosophy of mind and phenomenology* (pp. 48–66). Routledge.