

EXPLAINING INSERTED THOUGHTS
Matthew Parrott
St. Catherine's College, Oxford
(DRAFT: Forthcoming in *Philosophy, Psychiatry, and Psychology*)

It seems to be impossible for a person to have introspective access to thoughts that are not her own (Campbell, 1999). Yet, although first-personal conscious awareness of a particular thought is normally sufficient for being its owner, some schizophrenic subjects report being conscious of thoughts that are not theirs. This suggests that, contrary to philosophical orthodoxy, thought ownership is not a necessary condition for consciously experiencing a thought. Since what schizophrenics report is thus rather difficult to reconcile with standard philosophical conceptions of conscious thought, it would be good to have a clearer picture of precisely how experiences of thought insertion differ from those of ordinary thinking. Developing such a picture is the aim of Patrizia Pedrini's essay.

1. Dispositions and Agency

One of Pedrini's most valuable insights is that ordinary thinking is intimately tied to cognitive agency. Correspondingly, she thinks that experiences of thought insertion arise because of a certain type of agential failure. Although appeals to agency are familiar in discussions of thought insertion, the relevant failures have traditionally been conceived of primarily in volitional terms, either as some impairment in a subject's ability to generate her thoughts or as some missing sense of agency normally associated with this kind of act. Yet, as several theorists have pointed out, many ordinary thoughts, such as passing thoughts or free associations, also lack these features. So, if cognitive agency is uniquely impaired in cases of thought insertion, it must be in some other way.

Pedrini's conception of agency is importantly different. Rather than characterizing it in terms of acts or volitions, she offers a dispositional conception and develops her account of thought insertion along these lines. According to Pedrini, we are fundamentally cognitive agents and being such is a "condition of possibility of holding attitudes held to be our own." (pg. 9) More specifically, she claims that one must be a cognitive agent in order to experience a particular thought *as* one's own. This, however, does not require one to have actively formed the thought; rather it requires only that a person "be willing to take responsibility" for a thought and to actually do so "when (and if) the appropriate occasion arises." (pg. 9) Thus, cognitive agency as it figures in Pedrini's account should be understood as the disposition to "take responsibility" for one's thoughts.

But what does this involve? Among contemporary philosophers it is common to think a subject takes responsibility by endorsing her thoughts or judging them to be supported by reasons (Moran, 2001). This line of thinking is embodied in what Pedrini calls the Endorsement Model. According to endorsement theorists like Lisa Bortolotti (2010), a normal subject experiences a thought as her own because she has the capacity to endorse it with reasons (for similar views see Fernandez (2010) and Pickard (2010)). Pedrini's central criticism of this model strikes me as correct. Schizophrenics typically experience thought insertion with respect to a variety of attitudes, not simply rational ones. Thus, they are sometimes alienated from thought contents that they merely entertain, fancy, or attend to. It should be clear, however, that for modes of thinking like these the notion of *rational* endorsement is completely out of place. Indeed, it is hard to see what it could mean to judge that my entertaining a thought about the beach is supported by reasons. How could merely entertaining such an idea have or lack justification? At best, the Endorsement Model is

applicable to rational attitudes, like beliefs or desires, but not to the kinds of thoughts that, as Pedrini says, "do not have rational provenance" (pg. 10).

Is there some way to understand the notion of "taking responsibility" more broadly? If we look to social interactions, it seems that, in addition to citing reasons for our thoughts, we often cite causes. For example, if I am entertaining a thought about the beach and you were to ask why, I might reply that I am especially tired. In such a case, my fatigue would be cited as the cause of my thinking, not a reason. Since these sorts of causal explanations are as prevalent as rational ones, Pedrini proposes that we "take responsibility" for our thoughts in virtue of being disposed to engage in what she calls the "cause/reason-giving practice". We are responsible in the sense that we are inclined to explain what we are thinking, sometimes by giving reasons and sometimes by quoting "the causes of our attitudes, instead of their reasons" (pg. 10).

There is much about Pedrini's general picture of cognition that I find illuminating. Her dispositional conception allows us to see how agency can be connected to those parts of our mental lives that were not, or could not, be formed as the outcome of deliberation. Additionally, by emphasizing our dispositions to give causal, non-reason giving, explanations, Pedrini demonstrates how we can be agents with respect to non-rational attitudes. This basic framework is extremely appealing and I think it recommends novel approaches for thinking about several philosophical problems. Nevertheless, I am not fully convinced by Pedrini's use of it to explain thought insertion.

2. Explaining Inserted Thoughts

Pedrini wants to explain thought insertion on the basis of impaired agency, where this is understood in terms of dispositions to give rational or causal explanations for thoughts. She

writes, "if the occurrence of [an] attitude is not accompanied by the dispositions to act upon it in the 'cause/reason-giving practice' on appropriate occasions, the attitude is experienced as alien." (pg. 13) So, on her view, if a subject is not disposed to explain a particular thought, she will experience it as alien.

Prima facie, it is not implausible to think subjects have a *general* disposition to participate in the cause/reason-giving practice. This is typically how we envision dispositions to engage in practices that range over infinite domains. For instance, most of us are disposed to calculate sums whenever someone asks us to add two numbers. Most plausibly, we have a general disposition to calculate sums rather than a massive aggregation of distinct dispositions; i.e., a disposition to add 1+1, one to add 1+2, one to add 1+3, etc. Yet it seems this is not how Pedrini thinks of explanatory dispositions. She often speaks of a subject failing to "acquire" a disposition to engage in the cause/reason-giving practice with respect to a specific "piece of thinking" and this suggests that she believes subjects normally have a conglomeration of discrete explanatory dispositions, one for each thought.

In that case, however, it is much less clear whether ordinary subjects really do acquire this type of disposition for every single thought. Indeed, it seems to me that I sometimes have passing thoughts or daydreams that I am not at all disposed to explain; however, I am not thereby alienated from them. Pedrini will want to insist that, unlike the schizophrenic, I *do* have the relevant disposition but simply fail to manifest it on certain occasions. But, if explanatory dispositions are acquired on a case-by-case basis, it is hard to see what grounds she could have for this insistence. If you were to ask me to add 87 and 49 and I were not inclined to answer, it might be reasonable to say I am nevertheless disposed to calculate the sum; but this is because it is independently plausible that I am generally disposed to calculate sums. Without that sort of general disposition, it is hard to see what reason there would be

to maintain that I have the disposition to calculate this particular sum. Similarly, if we follow Pedrini in thinking that a discrete disposition is connected to each episode of thinking, it will be nearly impossible to distinguish cases where a disposition fails to manifest itself from cases where it is simply absent. Pedrini needs some criteria for making this distinction, especially if the latter are supposed to explain experiences of alienation.

My larger worry with Pedrini's account is that it looks to me like schizophrenic subjects *are* disposed to give causal explanations for their thoughts. Since they generally explain thoughts like we do, they appear to understand the cause/reason-giving practice perfectly well. Thus, if we look at firsthand reports, it is natural to think that schizophrenics are offering causal explanations for inserted thoughts. Subjects typically say things like "the thought was given to me electrically" (Mullins and Spence 2003, pg. 295) or "the houses had put them in my head" (Saks 2007, pg. 27). These are causal explanations. So why is a report of thought insertion not a manifestation of the disposition to participate in the cause/reason-giving practice, a disposition Pedrini claims is sufficient for experiencing a thought as one's own?

I think Pedrini attempts to address this question in the following passage:

The alienated thought lacks the agentive dispositions to enter the practice non deliriously, but then the subject, who attempts to make sense of her abnormal experience and whose global agentive capacity in thought are not impaired, makes an extreme move to enter into the practice: she introduces among the legitimate causes that explain why she has a certain perplexingly alien thought the possibility that another person can actually be thinking that thought in her place." (pg. 18)

This passage is perplexing. Pedrini begins by claiming the subject of an "alienated thought" lacks the relevant "agentive disposition" but she then says the subject has an unimpaired "global agentive capacity to enter into the practice". However, if the subject is "globally" disposed to explain her thoughts, then she is disposed to participate in the practice *simpliciter*. Any "local" disposition is redundant.

Perhaps Pedrini could appeal to the notion of "appropriateness". She might argue that a delusional explanation is inappropriate given the norms of the cause/reason-giving practice, so much so that offering one amounts to disengaging from that practice. But why do these explanations disqualify someone from making a move in the cause/reason-giving game? They are obviously highly irregular and they may even violate some kind of norms, for example epistemic norms governing the selection of a candidate hypothesis. Yet that doesn't unambiguously demonstrate the absence of a disposition to participate in the practice. Malapropisms, for example, show us how someone can be disposed to speak a natural language while nevertheless violating a norm on occasion. Is a delusional explanation that invokes inserted thoughts significantly different in some way? One might be tempted to think the key difference has to do with the fact that schizophrenics cite external causes rather than internal ones but that would be difficult to reconcile with what Pedrini says about the kinds of causal explanations we tend to give for our sensations or brute desires, which we obviously do not experience as alien.

Furthermore, offering bizarre causal explanations is characteristic of most schizophrenic delusions. For instance, one schizophrenic individual reported thinking he was a robot and that strangers nearly always recognized him. His explanation for this thought was that he was "in fact the subject of some kind of experiment designed to cure me from being so predictable" (Timlett 2013, pg. 245). Another claimed he felt distress because his

neighbors were "broadcasting their voices using black magic" (Freeman, et. al. 2004). It is typical for delusional explanations to be highly abnormal and idiosyncratic like these but they nonetheless exhibit the structure of causal explanation. Notice, however, that on Pedrini's view these explanations would have to be consistent with participation in the cause/reason-giving practice because their subjects do not experience alienation. But if explanations like these can manifest, or at least be compatible with, a disposition to engage in that practice, why would explanations invoking inserted thoughts be disqualified? It is hard to imagine an answer to this question that is not *ad hoc*.

Throughout her essay Pedrini urges us to concentrate on what is lost in experiences of thought insertion rather than on what she calls the "delirious" attributions associated with those experiences. But, I think it is rather likely that the two are closely connected. A substantive characterization of thought insertion might focus more on the relationship between the way in which a schizophrenic subject tries to describe her experiences and what those experiences are subjectively like. Perhaps an experience of thought insertion consists of some kind of impaired introspective access to one's own thoughts and a subject's delusional explanation is her attempt to characterize the nature of this impairment. Although this not the forum to pursue this line of thinking, it offers a way in which the phenomenon could be broadly consistent with our philosophical intuitions about introspection. But, more importantly, approaching the phenomenon from this angle could result in something Pedrini rightly values, a better understanding of what it is like for someone to experience thought insertion.

References

Bortolotti, L. 2010: *Delusions and Other Irrational Beliefs*. Oxford: Oxford University Press.

- Campbell, J. 1999: Schizophrenia, the Space of Reasons, and Thinking as a Motor Process. *Monist* 82: 609-626.
- Fernandez, J. 2010: Thought Insertion and Self-Knowledge. *Mind and Language* 25: 66-88.
- Freeman, D., Garety, P, Kuipers, E., Bebbington, P., Fowler, D. and Dunn, G. 2004: Why do People with Delusions Fail to Choose More Realistic Explanations for their Experiences? An Empirical Investigation. *Journal of Consulting and Clinical Psychology* 72: 671-680.
- Moran, R. 2001: *Authority and Estrangement*. Princeton: Princeton University Press.
- Mullins, S. and Spence, S. A. 2003: Re-Examining Thought Insertion. *British Journal of Psychiatry* 182: 293-298.
- Pedrini, Patrizia. forthcoming: Rescuing the "Loss-of-Agency" Account of Thought Insertion. *Philosophy, Psychiatry and Psychology*.
- Pickard, H. Schizophrenia and the Epistemology of Self-Knowledge. *European Journal of Philosophy* 6: 55-74.
- Saks, E. 2007: *The Center Cannot Hold*. London: Virago.
- Timlett, A. 2013: Controlling Bizarre Delusions. *Schizophrenia Bulletin* 39: 244-246.