Reply to John Campbell

Annalisa Coliva

John Campbell’s reply to my paper aims at re-establishing the point that there are two strands to our notion of ownership of a thought. There are two ways of cashing out this idea. First, one could say that A is the owner of a thought iff both the following two independent conditions obtain: (1) X is introspectively aware of a token thought and (2) X is the person who formed that token thought. Secondly, one may hold that there are two different and independent notions of ownership of a thought, call it O₁ and O₂, corresponding to (1) and (2) respectively. For brevity, I will refer only to the first interpretation. But what I will be saying will apply, mutatis mutandis, to both.

On this view, thought insertion would be a case in which someone thinks that (1) is satisfied, while (2) isn’t and, therefore, denies that the thought she is introspectively aware of is her own. Campbell's explanation of thought insertion is quite clear: the subject has prima facie reasons to think that she is not the person who formed that token thought – maybe because she does not experience that thought as formed by herself – and, therefore, has prima facie reasons to deny that she is its owner, but she is mistaken in identifying the producer of the thought and what she says is false, yet reasonable. Notice, however, an important consequence of Campbell's model. If (1) and (2) are independent conditions then the one could obtain without the other. Hence, it must be conceivable that one is introspectively aware of a thought that one has not produced. Indeed, cases of multiple personality, if taken literally, might be taken as examples of this kind of situation: person A and person B inhabit the same body, A can have access to B’s thoughts, yet B remains their producer. So A could actually say something like “I’m thinking (i.e. I am immediately aware) that p, but this is not my thought (it is not the thought I produced), it’s B’s (the thought that B produced)” and what she would be saying would in fact be true.

Yet, a simple-minded reaction one may have towards this line of explanation is this. If the subject is either saying something false, but reasonable, or, indeed, something altogether true, when she says “I’m thinking ‘Kill God’, but this isn’t my thought, it’s Chris’”, 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?

Hence, the whole dispute between Campbell and me hinges on whether it is logically possible for one to be introspectively aware of a thought and yet not be its producer. Campbell claims that it is, whereas I deny that this is a genuine possibility. He holds that these are two independent conditions and that the one can obtain without the other; whereas I hold that they are not independent and that they stand or fall together. Campbell is fully aware of being making a surprising claim because, as he points out, "[The two strands in the ordinary notion of ownership of a thought] do not ordinarily even seem to come apart" (Campbell XXXX, italics mine). Yet he offers in its support no more than an analogy between being the person who produced a certain signature and being the person who produced a given token thought. Let us look at it closer.

Consider the following situation: a person NN looks at a piece of paper with “NN” written on it and wonders whether she is the person who actually inscribed those letters on that piece of paper. After careful investigation she concludes that she is not the author of the signature and, therefore, says, “This signature isn’t mine”. On reflection, the example proposed by Campbell seems to exhibit the following features. (a) It is absolutely legitimate for the subject to wonder whether she is the author of that signature, although the name on the paper is her name and the signature looks pretty much like hers. Yet, it does not seem to be empirically possible that someone could be introspectively aware of someone else's token thoughts. (b) There are empirical investigations, which, in principle, could settle the issue. Still, we do not seem to have to find out that the person who produced a given token thought is the one who is introspectively aware of it. (c) If it turns out that, after all, that piece of paper has been signed by her, then we would say that she was mistaken and not ill. Finally, (d) if (c), then we might ask the person to revise, or correct her
judgement, but we wouldn’t think it appropriate that she undergo any treatment. Hence, if we take the analogy between being the author of a signature and of a thought at face value, it thus begins to struggle.

In general the point is that to be the person who produced a given signature is a causal (and strictly physical) notion and is thereby subject to ordinary empirical controls. So it ordinarily goes along with the empirical possibility that someone other than the subject may have forged the signature and that an empirical investigation can, in principle, determine who did it. By contrast, the break of the analogy on all fronts suggests that it is not an empirical possibility that someone other than the subject who is introspectively aware of a thought may be its producer, and that it is not open to empirical investigation whether things may be otherwise. So it is not clear – if we take the analogy at face value – that the notion of a producer of a given thought is a causal (and strictly physical) notion.

Thus, the analogy, if taken at face value, actually speaks against Campbell’s claim that conditions (1) and (2) are independent and in favor of identifying the notion of a producer of a given thought with the notion of the person who is introspectively aware of that thought. But if conditions (1) and (2) come down to the same condition, in the context of authorship with respect to a given token thought, then they stand or fall together.

If that is true, then we can easily explain why we cannot make rational sense of the logical possibility of being introspectively aware of a token thought and, at one and the same time, of rationally wondering whether one is the producer of the thought. Moreover, it becomes clear why we do not expect an empirical inquiry to settle this issue and we seem to know a priori that a given token thought has been produced by the subject who is introspectively aware of it. Furthermore, this would explain why we take the subject to be deluded and not just mistaken, if she were denying the authorship of her thought. For she would be denying of being the author of a thought which cannot possibly have been produced by anyone else but herself. Finally, it is because there is no rational explanation of what she is saying, that we take her report as a symptom not of a mistake but of mental illness.

But, of course, Campbell may claim that he is concerned only with the logical possibility that (1) and (2) come apart. In that case, however, the analogy of the signature ceases to be germane, for, as we have seen, that analogy, if taken at face value, has to do with empirical possibilities. So, Campbell has to re-interpret it as follows: (a) it is logically possible that a given token thought one is introspectively aware of has been produced by someone else. Hence, it is rational to wonder whether a given token thought one is introspectively aware of has been produced by oneself; (b) it is conceivable that things could turn out to be otherwise; (c) the subject, in our case, would make an error through misidentification relative to the producer of the thought. Even then it is still unclear why we would then try to cure the subject rather than point out to her that she should just revise her judgement, but I will not insist on that.

However, notice a crucial aspect of the dialectic: Campbell was supposed to give us reasons to believe that the notion of being the producer of a given thought is a causal one and is independent of the notion of being the person who is introspectively aware of a given thought. The analogy with the person who is the producer of a given signature was supposed to sustain that claim. In turn, that claim is the necessary precondition for arguing that it is logically possible that someone is introspectively aware of a given thought of which she is not the producer. Yet, in order for the analogy to work at all, Campbell has to re-interpret it in a way that presupposes the latter logical possibility. So, dialectically, the shuffle through the analogy proves merely to presuppose, rather than support, what Campbell wants to maintain.

It thus seems to me that Campbell has made no case by his analogy that there are two strands to our notion of ownership of a thought, rather than just one, namely what he calls the Introspectionist strand, i.e. condition (1). Therefore, this should help explain why I take it “as a datum from which [I] can argue that there is only one strand in the notion of the ownership of a thought”. (Campbell, XXXX).

Let me deepen the point. The analogy of the signature suggests that Campbell is here working with a notion of being the person who produced a given token thought which can be
glossed as follows: X is the producer of a given token thought iff X is the proximal physical cause of that token thought. Notice that, so far, there is nothing illegitimate. The problems arise when this notion is used to make rational sense of the patients' reports. For these reports are made by the subject on the basis of her own experience while having a thought. Yet a thought one is introspectively aware of does not seem to disclose its proximal physical cause. In fact, it is often remarked that thoughts do not carry a label with the name of their owner written on it. A fortiori, the mere having of a thought in one's stream of consciousness cannot manifest the individual to herself as its proximal physical cause. So, even if the notion of being the producer of a given thought so understood makes sense from a third person perspective, it is not a notion which, as such, can be used to make rational sense of the relevant reports from the first person perspective. For it is just not a notion the warranted application of which can be licensed by the subject's own experience. The interesting result is that if you decide to go along with Campbell, then, despite the original intentions, you would not be able to rationalize the patients' reports, for that would presuppose that the identification of the physical proximal cause of a thought could be warranted by introspection. So the alleged advantage of his account over mine, namely the fact that it could be used to make rational sense of the subjects' highly unusual experience seems spoiled.

Now, one may object that there is a further notion of being the producer of a given token thought that is not strictly physical, yet different from the notion of being the person who is introspectively aware of that thought. According to this notion, X is the producer of a given token thought iff X is the person who experiences that thought as integrated with the rest of her cognitive life – namely, with her long-term dispositions and desires – and as a thought on which, ceteris paribus, she would act. The advantages of this notion are clear: its application would be warranted by the subject's own experience and it would allow to make rational sense of the schizophrenic's report. But this is a problematic notion because it does not seem to have any direct bearing on the notion of ownership of a thought. For, if X does not experience T as integrated with the rest of her cognitive life, it does not follow at all, on our ordinary notion of ownership, that X is not the owner of T. After all, we do have whimsical thoughts that are not integrated with the rest of our cognitive lives but that do not thereby cease to be our own and to be recognized by us as such.

Of course, one may stipulate a different notion of ownership, whereby if X does not experience a thought as integrated with the rest of her cognitive life, then that thought would not be hers. But this would make the whole phenomenon of thought insertion slightly uninteresting from a philosophical point of view. For what would be the philosophical interest of saying that in order to describe certain psychiatric phenomena we need to invent or stipulate new categories? What would that teach us about our old ones? Yet the philosophical interest of this issue seemed precisely to be that it could teach us something important and otherwise unnoticed about our ordinary concept of ownership of a thought and that it might be used to impair a philosophical dogma such as logical immunity to error though misidentification of present-tense psychological self-ascriptions.

One may then suggest that, after all, there is a connection between the notion of being the producer of a given thought so characterized and the notion of being the owner of a given token thought. Consider the following analogy: X has just finished to write a poem. She reads it aloud and finds out that it does not fit at all with her original intentions, that is with the emotions and ideas she wanted to express. X could then say that the poem does not feel like being hers. She could actually go on to say that the poem sounds like Y's – another poet X happens to detest. Yet, by this analogy we would have identified a notion of feeling like being the owner of a given thought, but we would not have identified a different commonsensical notion of being the owner of a given thought. Moreover, all parties seem to agree that when the schizophrenic says "I'm thinking 'p', but this is not my thought, it is Y's" she does not mean the second part of the sentence metaphorically, or according to a secondary meaning of "being the owner of a given token thought". So the present notion of being the producer and, therefore, the owner of a given token thought, could not be applied in order to describe the patient's report, when taken at face value. Rather, it could only be used to re-interpret it.

Hence, in so far as we wish to make use of the notion of being the producer of a given token thought and, moreover, we require that its application should be warranted by the subject's
experience and could have a direct bearing on the issue of the ownership of that thought, we seem to have no other choice but thinking of the producer of a given token thought as the person who is introspectively aware of it.

Campbell then offers a tendentious reconstruction of the second part of my paper, which is worth considering in detail. I quote:

So the sequence of states [Coliva] is proposing is that (a) the subject finds herself with introspective knowledge of a thought, and (b) no sense of having been the producer of that thought, consequently (c) the subject is 'detached' with respect to that thought, or experiences it as 'a product of alien intentionality', and therefore (d) the subject concludes that she does not have introspective knowledge of the thought. Coliva does not put it quite this bluntly, but on her account of 'ownership', there is nothing else to enter at point (d). (Campbell XXXX)

What Campbell seems to be saying is that my account is incoherent because it first attributes the subject introspective knowledge of her thought and then denies that she has it. But in fact, what I am proposing is rather the following. (i) The subject finds herself with introspective knowledge of a thought, and (ii) with no sense of integration of that thought with the rest of her cognitive life. (iii) In extreme cases the subject experiences the thought as integrated with other beliefs and desires she is introspectively aware of, which, in turn, she does not experience as integrated with another body of beliefs and desires she has introspective knowledge of. (iv) The subject claims that the thought is not her own and (v) that it is someone else's. However, from (iv), on my account of ownership with respect to a given token thought, it follows (vi) that what she says commits her to the denial of having introspective knowledge of her thought. Hence, on my account, the subject herself is committed to a contradiction. Similarly, from the subject's avowal that she is introspectively aware of a given thought it follows, on my account, that she owns that thought. Hence, by denying to be its owner the subject commits herself to a contradiction. Yet this is rather different from saying that my account attributes ownership of a thought to X and then denies that X possesses that thought.

Moreover, Campbell accuses me of making opaque why the sequence I have just described should lead to a denial of ownership of the thought on the subject's part. (Cf. Campbell, p. XXXX). What he may mean is that what I say fails to individuate a fully rational path which would justify X's report. But I actually do not mind the allegation of not giving an account that makes the subject's report fully rational. In fact, on my account, since (ii) does not imply non-ownership, the passage from (ii) to (iv) involves a non sequitur. A fortiori the passage from (iii) to (v) is not warranted.

Let me make one final observation on the different kinds of explanation of psychiatric phenomena one may give. Campbell insists that he is trying to respect and provide a model of Frith's subpersonal account of thought insertion and that this kind of account belongs to "a wave of analyses in cognitive neuropsychiatry in which delusions are explained as broadly rational responses to highly unusual experiences" (Campbell, p. XXXX). But Campbell also attributes to me the view that my analysis of thought insertion aims at respecting Frith's subpersonal account. (Cf. Campbell, p. XXXX). Yet, nowhere in the paper do I make such a claim.¹⁰

There is, however, a more general point, which is worth noticing. On my account what the schizophrenic says implies a contradiction and it is arrived at through non rationally warranted movements of thought. On Campbell's account it does not imply a contradiction and, allegedly, it is arrived at through fully rational routes. Therefore, on my view there is no account which can rationalize what the patient is saying, just a reconstruction of a path that, however irrational it may be, may lead to a claim which implies a contradiction. What seems to me to be the interesting outcome of our debate, which may be of some interest to psychiatrists, is this. Of course, if you think that an explanation of psychiatric phenomena must give reconstructions that allow to see certain responses to unusual experiences as fully rational, then, in the case of thought insertion, you
have to eliminate the contradiction and see the reasoning behind the relevant reports as perfectly rational, although both these tasks may prove rather difficult. Yet, to take the opposite route, namely to acknowledge the contradiction and the non-rationally warranted paths which can lead to those reports does not by itself prevent one from providing some intelligible reconstruction of what is going on with the subject, its irrationality notwithstanding. Nor does this prevent one from studying the subpersonal mechanisms whose malfunctioning can cause, for instance, the occurrence, at the personal level, of thoughts that one finds at odds with other beliefs and desires of hers.

On this latter view, the phenomenon of thought insertion and the attempts at explaining it are relevantly comparable to Escher's drawings and their interpretation. As the presence in those drawings of contradictory images does not stand in the way of providing an understanding of them and the very fact that motivates the enterprise of understanding is the presence of non-eliminable contradictions, similarly the presence of contradictions in the schizophrenics' reports does not stand in the way of providing some understanding of them and indeed it may rather motivate the very enterprise of giving explanations in cognitive psychiatry.

1 I would like to thank Marco Panza, Gianfranco Soldati and Crispin Wright for their comments and criticisms to earlier versions of this reply.

2 I hope that by trying to be precise about what it means to say that there are two strands to our ordinary notion of ownership of a thought I am not doing any violence to Campbell's intuitions.

3 I take it that if (1) and (2) are such that the conjunction of one of them with the negation of the other is not a contradiction, as Campbell maintains (p. XXXX), this means that they are independent conditions.

4 By contrast, on the latter view, thought insertion would be a case in which someone owns (O1) a thought and rightly avows that, but thinks that (2) is not met and, therefore, says that she does not own (O2) that thought.

5 Notice, however, that we should still explain how the subject passes from the denial of the ownership of a token thought to the positive attribution of that thought to someone else. And, in particular, to the attribution of that token thought to someone who, at least in some cases of thought insertion, is not another person inhabiting the same body, but an altogether different (mental and) bodily entity.

6 Campbell himself insists that the phenomenon of thought insertion is an illusion, while, at the same time, he wants to explain it as a case of error through misidentification. (Cf. Campbell p. XXXX). A rough way of putting the difference between illusions and mistakes is to say that for the former we can only discern causes, while for the latter we can discern reasons, which should also be appraised by the subject.

7 Unless one believes that there are genuine cases of telepathy.

8 This may well not be true for sensations and perceptions, but it is difficult to dispute in the case of thoughts.

9 This, however, does not stand in the way of saying that a given thought is mine or that it is thought by me. For, as Peacocke would have it, the use of the first person is licensed by the evidence that the person who is introspectively aware of that thought is F.

10 We could call this kind of experience "sense of integration" as opposed to "sense of ownership", which is the label I actually use in the paper and that is often used in the literature. As a matter of fact, I do explain the meaning of that latter label in a way which makes it obvious to any sympathetic reader that by it I could not possibly mean a sense of being the person who is introspectively aware of a given thought. (Coliva XXXX).

11 The only claim I do make is that the loss of a sense of integration with respect to one's thoughts may be explained either at a personal or at a subpersonal level, but I do not commit myself to either view (Coliva, p. XXXX, and fn. 19).

12 For a criticism of Frith's account of thought insertion, cf. Lynn Stephens and Graham 2000: 140-144.