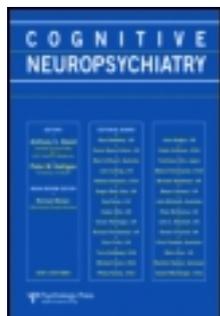


This article was downloaded by: [Glenn Carruthers]

On: 26 February 2012, At: 16:46

Publisher: Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Cognitive Neuropsychiatry

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pcnp20>

### A metacognitive model of the sense of agency over thoughts

Glenn Carruthers<sup>a b c</sup>

<sup>a</sup> Macquarie Centre for Cognitive Science, Macquarie University, Australia

<sup>b</sup> Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

<sup>c</sup> Discipline of Philosophy, University of Adelaide, Australia

Available online: 09 Dec 2011

To cite this article: Glenn Carruthers (2011): A metacognitive model of the sense of agency over thoughts, *Cognitive Neuropsychiatry*, DOI:10.1080/13546805.2011.627275

To link to this article: <http://dx.doi.org/10.1080/13546805.2011.627275>



PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## A metacognitive model of the sense of agency over thoughts

Glenn Carruthers<sup>1,2,3</sup>

<sup>1</sup>Macquarie Centre for Cognitive Science, Macquarie University, Australia

<sup>2</sup>Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

<sup>3</sup>Discipline of Philosophy, University of Adelaide, Australia

*Introduction.* The sense of agency over thoughts is the experience of oneself *qua* agent of mental action. Those suffering certain psychotic symptoms are thought to have a deficient sense of agency. Here I seek to explain this sense of agency in terms of metacognition.

*Method.* I start with the proposal that the sense of agency is elicited by metacognitive monitoring representations that are used in the intentional inhibition of thoughts. I apply this model to verbal hallucinations and the like and examine the plausibility of this model explaining deficits associated with these symptoms.

*Results.* By tying the sense of agency to metacognitive inhibition I propose that the loss of a sense of agency in certain psychotic symptoms is accompanied by a particular deficit in the patient's ability to control their own thinking. This is consistent with the experiences of those at high risk of developing hallucinations, who report more intrusive thoughts than controls. The model I present is able to explain why those at risk of developing verbal hallucinations and those suffering from verbal hallucinations have deficits in the intentional inhibition of thought. I defend this account from a possible objection by distinguishing the form of the intentional inhibition deficit displayed by those suffering verbal hallucination from that displayed by those suffering from orbital-frontal cortex lesions and posttraumatic stress disorder.

*Conclusion.* A plausible hypothesis is that the sense of agency over thoughts is elicited by the metacognitive monitoring representation used to intentionally inhibit

---

Correspondence should be addressed to Glenn Carruthers, Macquarie Centre for Cognitive Science, C5C, Macquarie University, Balaclava Road, NSW 2109, Australia. E-mail: gcarruth@maccs.mq.edu.au

Parts of this project were undertaken towards a PhD at Macquarie University. My thanks go John Sutton and Max Coltheart for supervising the project. I also wish to thank Georgie Paulik and Laura May Botterill for reading early drafts. I am indebted to the editor Peter Halligan and the reviewers for *Cognitive Neuropsychiatry* whose challenges made this a much better paper.

---

© 2011 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business  
<http://www.psypress.com/cogneuropsychiatry> <http://dx.doi.org/10.1080/13546805.2011.627275>

## 2 CARRUTHERS

thoughts. The deficit in the sense of agency over thoughts associated with certain psychotic symptoms could be explained by a failure to properly metacognitively monitor certain thought processes.

**Keywords:** Metacognition; Self-consciousness; Sense of agency over thought; Verbal hallucinations.

### INTRODUCTION

In this paper I provide a model of the sense of agency over thoughts. This is the sense that one is the thing that does the thinking in question, that one is the agent of mental action. This feeling is generally taken to be deficient in a variety of symptoms of schizophrenia, namely verbal hallucinations (especially voices commenting and voices arguing), delusions of thought insertion, delusions of thought withdrawal, made impulses, and made emotions. Typically these symptoms are taken to arise (in part) from the patient misrepresenting themselves as not the agent behind thoughts they control (see, e.g., Campbell, 1999; Frith, 1992, p. 66; Graham & Stephens, 1994; Mellor, 1970, p. 17; Young, 2008). Commonly it is supposed that the patient fails to self-attribute a thought or piece of verbal imagery because they lack a sense of being the agent of the thought. In contrast they appear to know that the thought occurs in their mind, or that it is they who hear the voice. It appears that their sense of being the subject of a thought (sometimes called a sense of ownership<sup>1</sup>) is intact. As such a deficit in the sense of subjectivity (ownership) is intact and a deficit in this sense is not the reason patients fail to self-attribute thoughts. However, it may also be the case that these patients veridically represent themselves as not the agent behind thoughts they don't control. Here I suggest that it is not easy to differentiate these options as a problem with eliciting a sense of agency and a problem with inhibiting unwanted thoughts (which is one form of control over thinking) can have a common cause. Despite this, a consideration of these symptoms is still a good place to start when attempting to model the sense of agency. I will begin with a quick summary of some basic findings in metacognitive research which will allow for the presentation of the model I propose here. I will then present the model in general and apply it to explaining the deficits in the sense of agency seen in certain psychotic symptoms. Following this I will provide some preliminary evidence for the

---

<sup>1</sup> However, one needs to be careful here; the term "ownership" sometimes refers to a feeling of being the subject of a thought (Gallagher, 2000) and sometimes to an attribution judgement (Campbell, 2002). If one uses the prior sense it is thought to be deficient in these symptoms. If one uses the latter sense then it is more controversial to claim a deficit (Bortolotti & Broome, 2009). As it is only the sense of agency under discussion here, I will not pursue this further.

model by showing that there are difficulties with intentional inhibition correlated with sense of agency deficits amongst those either suffering from or at high risk of developing verbal hallucinations. To close I defend the model I present from a possible objection by distinguishing the form of the intentional inhibition deficit displayed by those suffering verbal hallucination from that displayed by those suffering from orbital-frontal cortex (OFC) lesions and posttraumatic stress disorder (PTSD).

## METACOGNITION

In order to develop a model of the sense of agency over thoughts in terms of certain forms of metacognitive monitoring I begin with a brief discussion of some of the basic findings in the study of metacognition, which I use to argue for a distinction between “object-level cognition” and metacognition (Nelson & Narens, 1990, pp. 125–126). Furthermore, I use this discussion to distinguish two types of representation which are required for metacognition, I call these “metacognitive beliefs” and “monitoring representations”.

Metacognition is the representation and regulation or control of cognitive processes (Flavell, 1979; Nelson & Narens, 1990, p. 126). It is to be distinguished from theory of mind in that it does not involve a representation of mental states as representations, but rather a representation that mental processes are occurring and what can be achieved via those processes (Proust, 2003). Metacognition is usually operationalised in terms of people’s insight into their cognitive abilities. For example, if preschool children and primary school children are given a list of items to memorise and asked to indicate when they can remember the entire list, primary school children can accurately indicate when they have memorised the list, whereas preschool children cannot. This is taken to indicate that the older children have better insight into their mental abilities, thus they have more developed metacognitive skills (Flavell, 1979, p. 906).

To get a grasp on metacognition, it is helpful to distinguish “object-level” cognition from metacognition (Nelson, 1996, p. 105; Nelson & Narens, 1990). The object-level processes are those which are taken as the object of representation of the metacognitive representations. These are typical cognitive processes such as motor control, perception, learning, and belief formation (Nelson, 1996, p. 106). Metacognition is the representation and regulation of these object-level processes.

There are (at least) two types of metacognitive representations. The first I shall call monitoring representations. These represent specific object-level processes as they occur. Examples of these are the “feeling of knowing” and “judgement of learning”, which will be discussed later (Jost, Kruglanski, & Nelson, 1998, p. 140).

The second type of metacognitive representation is relatively more stable and takes the form of beliefs or theories about how object-level processes work. I shall call these metacognitive beliefs, although they have also been called “metacognitive knowledge” (Flavell, 1979, p. 906). These represent one’s cognitive processes in general, not a specific occurrence of a process. An example of such a belief is the belief that one is good at maths (Jost et al., 1998, p. 141). Other examples include confidence in one’s memory or beliefs about how often one has unwanted thoughts (Cartwright-Hatton & Wells, 1997). Such beliefs can also have a normative character, such as beliefs about the logical consistency of thoughts (Lobban, Haddock, Kinderman, & Wells, 2002).

This distinction has gone by many names. For example, Koriat (2000, p. 149) names monitoring representations “metacognitive feelings” and metacognitive beliefs “metacognitive judgements”. Later, correctly rejecting the idea that monitoring representations must be conscious, Koriat (2007, p. 344) names monitoring representations “experience based judgements” and metacognitive beliefs “theory based judgements”. I prefer the monitoring representation/metacognitive belief terminology as it doesn’t imply that monitoring representations are all about other conscious representations and it avoids the need to provide an analysis of the term “judgement” or “knowledge”.

The key difference between metacognitive beliefs and monitoring representations is what they represent. Monitoring representations are about a specific occurrence of a process; in contrast, metacognitive beliefs are about the subject’s cognition more generally. To tell metacognitive beliefs apart from monitoring representations it may be helpful to consider how each can be operationalised. Metacognitive representations are typically operationalised via people’s reports of their own cognitive performance (Jost et al., 1998, p. 139; Nelson, 1996, p. 106). This can be done in different ways depending on whether or not the target is a metacognitive belief or a conscious monitoring representation. To measure the more stable metacognitive beliefs, questionnaires that ask subjects to report beliefs about what their abilities are or what they ought to be are appropriate. An example of such a questionnaire is the Metacognitive Questionnaire (MCQ), which contains items such as “I have little confidence in my memory for names and faces” and “I pay close attention to the way my mind works” (Cangas, Errasti, Garcia-Montes, Alvarez, & Ruiz, 2006 p. 489; Cartwright-Hatton & Wells, 1997). To get to conscious monitoring representations, subjects are asked for reports of their performance on a specific task. This can be done by, say, asking the subject to indicate when they are satisfied that they have learned some material, or more indirectly by methods such as postdecision wagering (Seth, 2008) or other measures of confidence. Of course, unconscious monitoring representations cannot be operationalised

by direct report. Instead, they are typically measured by the effects they are thought to have on object-level cognition. An example of this is discussed later when I consider the role of metacognition in regulating object-level processes.

The metacognition/object-level distinction is supported by the fact that they can be altered independently of one another. Consider the feeling of knowing. This is the sense that one was previously familiar with some information, but is not anymore (that one has heard it before) or that one would recognise the answer to a question were it given. It is a conscious monitoring representation and has as its content a particular attempt to remember. It is closely related to “tip of the tongue” phenomena (Nelson, 1996, p. 108). In one study subjects were presented with a series of general knowledge questions. After the question they were subliminally presented with the answer or a nonsense letter string. Those who were presented with the answer were more likely to answer correctly. That is, the object-level process was facilitated by the subliminal presentation of the answer. However, there was no difference between the two groups on their feeling of knowing judgements (Nelson, 1996, p. 108; see also Koriat, 2007, pp. 336–338, for a review of this dissociation). This suggests that it is possible to affect the object-level process of recall without affecting the metacognitive representations of those processes; hence, the processes are distinct.

Both types of metacognitive representation play a role in the regulation of object-level processes.<sup>2</sup> This is clear for metacognitive beliefs about mathematical skill. Women who believe that women are worse than men at mathematics do worse on mathematical tests than men. However, if the same women are told that on a particular test there is no gender difference regarding performance, they do as well as men on that test (Jost et al., 1998, p. 142; see also Koriat, 2007, pp. 331–336, for a review). It is clear that beliefs about mathematical ability can influence mathematical performance. I will leave metacognitive beliefs here, as the second type of metacognitive representation, that is monitoring representations, play a much greater role in the account of the sense of agency over thoughts offered later.

Monitoring representations, such as the feeling of knowing discussed earlier, also play a role in regulating and controlling object-level processes. Another type of monitoring representation that has been studied in some detail is the judgement of learning. Judgements of learning are judgements

---

<sup>2</sup>Note that there are other metacognitive representations that are also involved in the regulation of object-level processes. So-called “implementation intentions” are an example (see Gollwitzer & Schaal, 1998, for a comprehensive discussion). However, these representations play no role in the models of the sense of agency over thoughts being presented here, so I will leave these issues aside.

that one has learned some material, say some mathematical procedure or what happens in a story. Such judgements are used as a motivation to continue or stop learning (Nelson, 1996, p. 111). This may seem obvious, but it is significant in that it shows that monitoring representations as well as metacognitive beliefs play a role in the regulation of object-level processes.

These examples of monitoring representations are themselves all conscious representations. As these are the most widely discussed examples of monitoring representations, it is easy to suppose that all, or at least typical, monitoring is done by conscious monitoring representations. Shimamura (2000, p. 313), for example, suggests that “[m]etacognition refers to evaluation and control of one’s cognitive processes. In this way, metacognition often suggests conscious or volitional control of thoughts, memories, and actions.” At times when unconscious monitoring representations are discussed they can be viewed as significantly less important than conscious monitoring representations (e.g., Koriat & Levy-Sadot, 2000).

But, of course, not every representation involved in metacognition is itself conscious. Indeed there are unconscious monitoring representations. Reder and Schunn (1996), for example, suggest that when subjects have several possible strategies available to them to solve a problem they will tend to choose the strategy that has previously been successful, even when they are unaware of the different strategies and their associated rates of success. This suggests that the use of the strategies is unconsciously monitored. In one task subjects were asked to identify a series of multiplication equations as true or false, e.g.,  $8 \times 4 = 32$  is true but  $8 \times 4 = 33$  is false. There are of course several ways to check the equations that subjects might use. One rule which subjects did seem to use is termed by Reder and Schunn the “parity rule”. This rule states that the product of two even numbers is always even, if using this rule (or something similar) then equations of the form  $\text{Even} \times \text{Even} = \text{Odd}$  can be quickly identified as false, without determining the actual solution. Subjects were sensitive to this fact, being five times faster in identifying false equations when 80% of false equations were of this form than when 20% were (p. 67). However, subjects were unaware of the different proportions of equations of this form on each run and many were unaware of the parity rule or any equivalent (p. 67). This suggests that the monitoring representations which lead to the determination that the more efficient strategy (applying the rule rather than determining the actual product) is a successful strategy were unconscious.

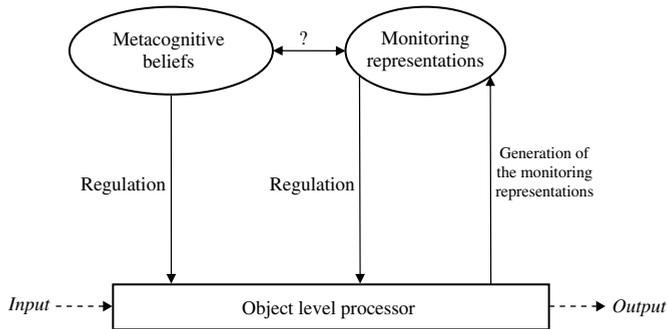
This discussion of metacognitive representations suggests the following model. There are object-level cognitive processes that are the represented objects of metacognitive representations. There are two distinct types of

metacognitive representation, these I have called monitoring representations and metacognitive beliefs. Both of these are involved in the regulation of object-level processes and may themselves be conscious or unconscious. They influence when object-level processes occur and also specific performance of the process. The relationship between the two types of metacognitive representation has been left open, this is indicated by the “?”. This is summarised in Figure 1.

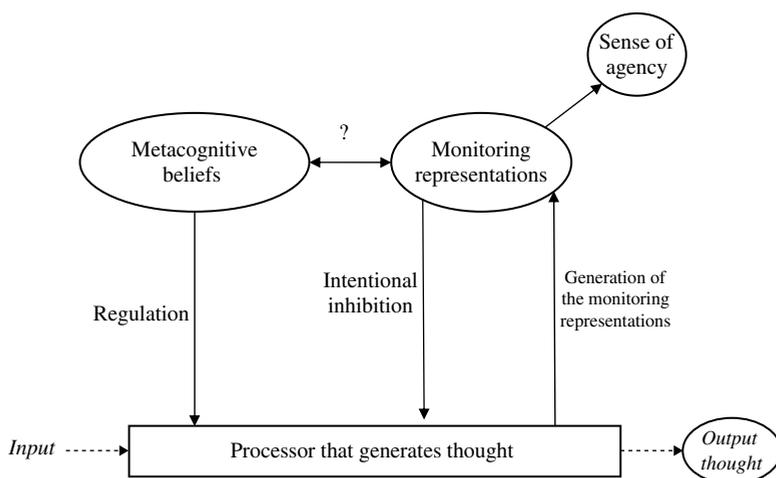
## A METACOGNITIVE MODEL OF THE SENSE OF AGENCY

On the general model of metacognition introduced above certain object-level processes are represented by monitoring representations and regulated by these representations as well as by beliefs about the processes (Figure 1). Here I wish to propose that the sense of agency over a thought is elicited by certain monitoring representations of, but not metacognitive beliefs about, particular object-level processes. I propose that the object-level process involved is any process that counts as generating a thought and that the metacognitive representations involved are those involved in the intentional inhibition of unwanted thoughts. I will apply this model to the sense of agency deficits in verbal hallucinations. My task will be to show how the sense of agency could fail to be elicited by this computational system in this case. The model is outlined in Figure 2.

Figure 2 shows an object-level process that produces a thought (e.g., eliciting a memory). This process can be intentionally, i.e., deliberately,



**Figure 1.** Metacognition is the representation and regulation of some object level cognitive process (shown here as a generic processor box). Two types of metacognitive representations are shown, these are monitoring representations and metacognitive beliefs. These are particular representing vehicles delineated by their representational content. As with all diagrams here, these vehicles are drawn as ellipses. Solid arrows indicate the computations particular representing vehicles enter into and dashed arrows indicate the transfer of information (without it being altered). The “?” indicates that the relationship between the two types of metacognitive representation has been left open.



**Figure 2.** A model on which the sense of agency over a thought is elicited by some monitoring representation of a process of generating a thought, where the monitoring representation is one that can be used in the intentional inhibition of thought.

inhibited using metacognitive beliefs (such as the belief that one has a bad memory which would lead to fewer attempts to retrieve the memory) as well as by monitoring representations (such as the feeling of not knowing which would lead to attempts to retrieve a specific memory being inhibited). Inhibition of a process is one of the ways in which the metacognitive processes can regulate or control the object level. I suggest that it is those monitoring representations which are potentially used to intentionally inhibit the process of generating a thought that cause the sense of agency to be elicited. In this paper, as I am focused on the relation between these monitoring representations and the sense of agency, I will not provide further details as to how the process of inhibition works. Think of this as an additional functional characterisation of the representation that elicits the sense of agency.

The sense of agency over one's thoughts is often thought to be deficient in some delusions and hallucinations associated with schizophrenia; most prominently with the delusion of thought insertion and verbal hallucinations. Here there is a problem. Are these patients: (1) misrepresenting themselves as not the agent behind thoughts they control, or (2) veridically representing themselves as not the agent behind thoughts they don't control? That is, do patients with these delusions and hallucinations fail to represent themselves as being in control of properly voluntary thought processes? Or

are they (more or less) accurately reporting that thoughts are appearing to them via a process they don't actually have control over?<sup>3</sup>

This question deserves more attention with regards to verbal hallucinations, thought insertion and the like. It is often taken as a working assumption that these patients' abilities to control their thinking are largely intact, with just their ability to represent who the agent of a thought is being impaired. However, as these options have different implications for our understanding of how verbal hallucinations arise, it is worth considering both alternatives. In particular it may be the case that attempts to model verbal hallucinations in terms of a deficit in the sense of agency alone gloss over other significant cognitive problems that may also contribute to the emergence of hallucinations (Paulik, Badcock, & Maybery, 2008, p. 34).

If the deficit in these symptoms is a metacognitive deficit, specifically a deficit in monitoring representations, then I would predict that, rather than simply losing the ability to represent themselves as agents of mental actions; these patients also lose some ability to control their thinking. As, by hypothesis, the monitoring representations also play a role in the regulation of object-level processes, any deficit in the monitoring representation should also lead to a deficit in the regulation of the object-level process. If the object-level process in question is one of generating thoughts, then a deficit in monitoring representations of these processes will lead to deregulation of thought generation. Specifically, I suggest that, in order to deliberately inhibit a thought, the process generating that thought must be represented so that the specific thought can be selected for inhibition. A deficiency in this monitoring will prevent the subject from being able to select a thought for (potential) inhibition. Thereby, on this model those suffering from verbal

---

<sup>3</sup>The issue as to what the agent of a thought is a very interesting one, but need not be analysed in great detail to determine if a patient's sense of agency accurately represents their control over their thinking. Let me explain. The control of thought, we would expect, is likely due to a variety of dissociable mechanisms (one, perhaps, being the metacognitive intentional inhibition mechanism being discussed here). Now, a person would have control of their thinking in virtue of having these mechanisms as part of their make-up. We could thus say they are agents of thought in virtue of possessing all of these mechanisms. If a single one of these mechanisms were impaired then, barring compensation from other mechanisms, the person's agency/control over their thoughts would be impaired. For the person to have a normal sense of agency in such a case is to misrepresent the state of themselves. They are no longer in control of their thinking to the same degree. Similarly, if one lacks or has an abnormally weak sense of agency when all of these mechanisms are intact and functional, they misrepresent themselves, they represent themselves as not in control when, in fact, they are. Many theories of the deficit of the sense of agency in verbal hallucinations and the like assume that this latter possibility is true. However, an important implication of model I am proposing is that this lack or abnormally weak sense of agency in these symptoms accurately represents (in an obviously impoverished way) the patients reduced control over their thinking in virtue of a malfunctioning of one of the control mechanisms (the intentional inhibition mechanism).

hallucinations and the like are taken to have a deficit in controlling their thoughts as well as their sense of agency being absent.

That there is a deficit in patients' ability to control their thinking seems to be supported by the fact that those with a high hallucination predisposition report more intrusive or unwanted thoughts than those with a low predisposition. Those at high risk of hallucinations are more likely to agree with statements such as "no matter how hard I try to concentrate, unrelated thoughts always creep into my mind" (item taken from Launey-Slade Hallucination Scale; Bentall & Slade, 1985).<sup>4</sup>

Two studies (Holmes & Steel, 2004; Marzillier & Steel, 2007) have examined intrusive thoughts in those at high risk of developing hallucinations. Both of these studies examined intrusive thoughts about traumatic events. This provides a way to control the nature of the intrusive thought and a guarantee that subjects attend to them. Intrusions are thus defined as:

unwanted and uncontrollable thoughts, images, memories and impulses that involve re-experiencing a traumatic event and are associated with high levels of distress.  
(Marzillier & Steel, 2007, p. 60)

Intrusions in this specific case, as well as more generally, are similar to verbal hallucinations except that they seem to come from oneself as opposed to an outside source (Paulik, Badcock, & Maybery, 2006, p. 1068). The studies of Holmes and Steel (2004) as well as Marzillier and Steel (2007) establish the same conclusion, so I will discuss just one of them here. Holmes and Steel attempted to induce this kind of intrusive thought in healthy individuals. This was done by showing subjects video of the aftermath of road accidents. Subjects kept diaries recording their intrusive thoughts about the video over the next week. Subjects were also given the Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE), one subscale of which, namely the unusual experiences subscale, is thought to measure a risk factor for the development of positive symptoms of schizophrenia including verbal hallucinations, thought insertion, and the like (Holmes & Steel, 2004 p. 28). It was found that those who had high scores on this scale tended to have significantly more intrusive thoughts about the video than those who had low scores (p. 31). So it seems that those at high risk of developing verbal hallucinations experience more intrusive thoughts than those at low risk. This, in turn, suggests that those at high risk of developing hallucinations have poorer control over their thoughts, thus favouring Option 2. This is consistent with the earlier model that predicts a link between the sense of agency deficits and the deregulation of thinking.

---

<sup>4</sup>My thanks go to Georgie Paulik for bringing my attention to this.

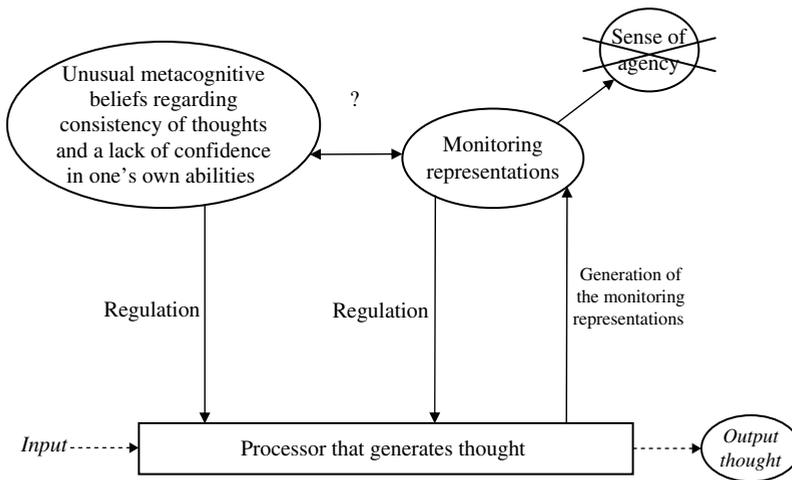
In the next section I will examine evidence consistent with the hypothesis that these intrusions come from a failure to intentionally inhibit thoughts. Unfortunately, these studies are all correlational. They show a correlation between sense of agency deficits, or vulnerability to such deficits, and intentional inhibition problems. Inferences regarding causation are not licensed by these studies alone. However, the results of these studies would be predicted by the above model, this fact does provide preliminary evidence for the hypothesis that sense of agency deficits are due to deficits with particular metacognitive representations.

According to this model the failure of the process of intentional inhibition is caused by a failure of metacognitive representation. In particular, I propose that it is a deficit in the formation of those monitoring representations that are about the processes that generate (but not those processes which regulate the generation of) thoughts. On the model I have presented here this would also explain why these subjects have a predisposition to not feel that they are the agent of their thoughts. If the sense of agency over thoughts is elicited by certain monitoring representations, then if these monitoring representations are degraded then necessarily the sense of agency will also be degraded.

Furthermore, if this model is correct then if the monitoring representations are degraded then the regulatory processes they are involved in will also be impaired. If the particular regulatory process affected is, as I argue later, one of intentional inhibition, this suggests that the specific monitoring representations that are implicated in hallucination predisposition are those that are involved in the intentional inhibition of object-level processes. I suggest that it is a problem with this process that explains the greater number of intrusive thoughts observed by Holmes and Steel (2004). This model of the hallucination predisposition is presented in Figure 3.

Figure 3 is a model of the problems with metacognition present in those who are at risk of experiencing verbal hallucinations. Here I argue that these subjects have a degraded ability to intentionally inhibit unwanted thoughts, degraded in the sense that sometimes it doesn't work. I examine evidence for this claim in subjects' difficulties with inhibiting currently irrelevant memories later. If this model is correct these subjects sometimes fail to inhibit memories due to their occasional failure to adequately monitor the process of eliciting the memory. Without this monitoring it is not possible to select a memory for inhibition. It is this occasional failure of monitoring that prevents irrelevant memories from being inhibited. It is also this impairment of monitoring that would account for these subjects' predisposition to hallucinate and explains why these subjects experience more intrusive thoughts.

This also suggests a model of the abnormal sense of agency in verbal hallucinations and the like. Perhaps verbal hallucinations and related

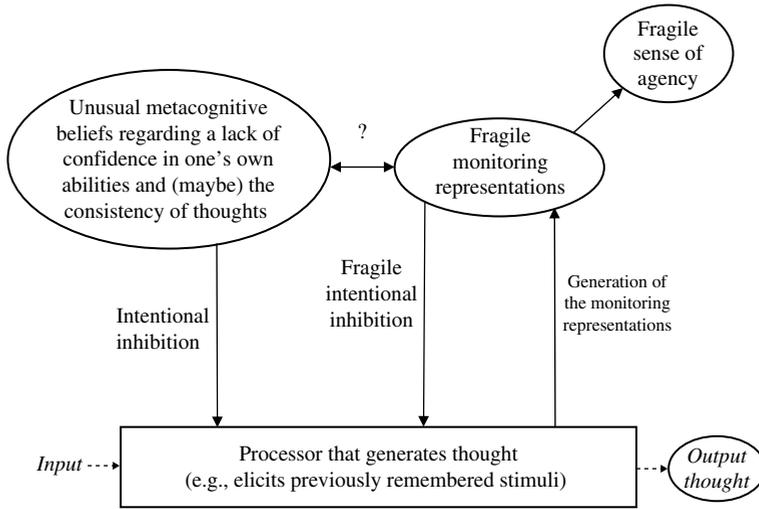


**Figure 3.** The first step towards providing a model of the metacognitive deficit in verbal hallucinations. Those with verbal hallucinations have unusual metacognitive beliefs and a deficit in their sense of agency.

delusions arise when the monitoring representations become truly *deficient* or absent, as opposed to merely impaired or degraded. This would leave a deficient sense of agency. On this model, however, it would not completely destroy the capacity for the intentional inhibition of thoughts as metacognitive beliefs as well as monitoring representations are thought to contribute to this process.<sup>5</sup>

This model can easily be extended to the abnormal sense of agency in delusions of thought insertion, thought withdrawal, made emotions, and made impulses. The difference between verbal hallucinations, thought insertion, made emotions, and made impulses is simply the object-level process that is not being monitored properly. Exactly which process is indicated by the phenomenology of each symptom? I suggest that in the case of verbal hallucinations it may be verbal imagery that is not monitored properly. Whereas, in the case of made emotions and made impulses, I suggest a deficit in monitoring the generation of emotions and impulses or intentions, respectively. In contrast, for thought insertion it may be thoughts more generally. Thought withdrawal, on the other hand, may involve a deficit in the formation of monitoring representations of the ending of a thought. This model is summed up in Figure 4.

<sup>5</sup>However, there is some reason to suspect that this could also be impaired by, e.g., a patient's unusual belief that their thoughts should be highly consistent with one another (see, e.g., Lobban et al., 2002).



**Figure 4.** A model of hallucination predisposition. A degraded monitoring representation causes a degraded sense of agency and problems with the intentional inhibition of thoughts.

Verbal hallucinations come in many varieties, three of which have special diagnostic significance for schizophrenia. The first of these is referred to as “audible thoughts” and involves the patient hearing their own inner speech being spoken aloud. This kind of hallucination may not involve a loss in the sense of agency, but could arise out of an unusual sense of subjectivity, e.g., the unusual experience that one is not the only person with access to one’s own thoughts. This is not a straightforward issue to resolve, so to avoid this complication any analysis of audible thoughts could be taken to be a separate project from an analysis for voices arguing or voices commenting. It may be problematic to treat all of those patients suffering from verbal hallucinations as having a similar deficit as is standard in the literature.

The remaining two kinds of verbal hallucination more obviously seem to arise from a deficit in the sense of agency. These are voices arguing, which involves two or more voices discussing or arguing about the patient (Mellor, 1970, p. 16), and voices commenting, in which a hallucinated voices comments on the patient’s actions as they occur (Mellor, 1970, p. 16).

In the model presented previously, I propose that those suffering from verbal hallucinations have problems with the representation of and their ability to intentionally inhibit unwanted verbal imagery or inner speech. These patients should have some ability to inhibit these unwanted thoughts

by using metacognitive beliefs about their thinking.<sup>6</sup> In this model the major problem in the regulation of their thoughts arises as they cannot form proper monitoring representations to use to inhibit unwanted thoughts as they occur. This deregulation can explain why these patients lack a crucial aspect of their control of their own thinking. Some thoughts cannot be dismissed. I will argue later that there are other ways in which inhibition can be deficient without an accompanying deficit in the sense of agency. Furthermore, the lack of proper monitoring of the processes of thinking means that there is nothing to elicit a sense of agency for these thoughts. Hence, these subjects lack a sense of agency for these intrusive thoughts.

In the model I present here, I wish to suggest that the sense of agency deficit arises for delusions of thought insertion, thought withdrawal, made impulses, and made emotions symptoms in a manner similar to that proposed earlier for verbal hallucinations. As already mentioned, what may differentiate these symptoms from each other and verbal hallucinations is the object-level process, which is deregulated due to the deficit in monitoring representations. As this kind of monitoring is supposed to be in place for any of these cognitive processes, this general model can account for the sense of agency over each type of thought. But, why is this kind of explanation plausible? Are there metacognitive deficits associated with these symptoms?

### INTENTIONAL INHIBITION DEFICITS ASSOCIATED WITH DEFICITS IN THE SENSE OF AGENCY

In order to develop and test this model I consider what metacognitive deficits are associated with verbal hallucinations and the like. Do these patients display problems with metacognitive representation? What about metacognitive regulation? Which object-level processes, if any, are disrupted by these problems?

Recent work by Paulik, Badcock, and Maybery (2007, 2008) suggests the hypothesis that sense of agency deficits are due to problems with particular metacognitive representations. Their data suggest that those predisposed to verbal hallucinations have difficulties with the intentional inhibition of previously relevant, but currently irrelevant, memories. If these results generalise to other kinds of thought, this will provide preliminary evidence for the claim that the sense of agency is related to the intentional inhibition of thoughts, as claimed earlier.

---

<sup>6</sup>However, these patients have somewhat unusual metacognitive beliefs, especially with regards to the consistency of their thoughts and their unusually low confidence in their abilities (Lobban et al., 2002). This may make inhibition more difficult still.

Paulik and colleagues (2007) assessed a group of healthy students using the Launey-Slade Hallucination Scale–Revised (LSHS-R). LSHS-R is a questionnaire that has been shown to measure predisposition to verbal hallucinations. They took two groups of students based on their scores on this questionnaire. A random sample of those who scored in the top 10 percentiles of the questionnaire scores made up the high hallucination predisposition group. A random sample of those from the bottom 10 percentiles of the questionnaire scores made up the low hallucination predisposition group (p. 460). These groups were then tested on the inhibition of currently irrelevant memories task.

The inhibition of currently irrelevant memories task is designed to measure subjects' ability to intentionally (i.e., deliberately) inhibit memories. In this task subjects are shown a series of pictures on a computer screen. The subjects' task is to indicate if they have previously seen a picture by pressing a "yes" or "no" button. Subjects run through three sequences of pictures, with the same pictures used in each sequence. The task is to indicate only when pictures are repeated within a sequence and ignore ("forget about") the pictures seen on the previous sequences (Paulik et al., 2007, p. 460). The task therefore requires subjects to intentionally inhibit memories of previous-sequence items (p. 460).

It was found that the low and high hallucination predisposition groups did not differ on the first sequence of pictures. This indicates that both groups were equally capable of recognising pictures. The high hallucination predisposition group made significantly more false recognitions on Runs 2 and 3 than the low hallucination predisposition group. This indicates the high hallucination predisposition group have impaired intentional inhibition. These differences were significant even after current anxiety<sup>7</sup> and IQ were controlled for (Paulik et al., 2007, p. 463). In addition to measuring the number of false recognitions, Paulik and colleagues also measured the time it took subjects to respond to each picture. On Runs 2 and 3 there was no difference between the groups in the times it took subjects to correctly identify a repeated picture. In contrast, the high hallucination predisposition group took longer to correctly identify a picture as not a repeat; again, this suggests difficulties with intentional inhibition. This held even after anxiety and IQ were controlled for (p. 464).

This suggests that those who have a high hallucination predisposition have difficulty with intentionally inhibiting certain thoughts, in this case previously relevant memories (Paulik et al., 2007, p. 465). In turn, this suggests a relationship between sense of agency problems and intentional inhibition problems as required by the model presented previously. However,

---

<sup>7</sup>This is important as some anxiety disorders are also associated with deficits in metacognition (Lobban et al., 2002).

this model does not merely suggest that this deficit occurs in those at risk of developing hallucinations. Rather, this model suggests that those suffering from verbal hallucinations and the like will show a deficit in their ability to intentionally inhibit thoughts.

This would seem to be supported by a study from Racsmány and colleagues (2008), who provide evidence that those suffering from schizophrenia have problems with intentionally inhibiting memories of words. Unfortunately, however, Racsmány and colleagues provided no analysis by symptom. As such it is not known if this problem arises for those suffering verbal hallucinations, delusions of thought insertion, and the like, or if it is a more general deficit associated with the entire syndrome of schizophrenia. Hence, this data cannot be taken as evidence for or against the model presented earlier.

That being said, there are other studies which do speak to the plausibility of this model. Badcock, Waters, Maybery, and Michie (2005) compared the performance of those suffering from schizophrenia both with and without verbal hallucinations to healthy controls on the inhibition of currently irrelevant memories task described earlier. Healthy controls were better than both patient groups in identifying repeated pictures, with both patients groups performing at the same level in this regard (p. 130). Furthermore, all three groups made the same amount of false recognitions on the first run. However, on the three subsequent runs those patients who were currently suffering from verbal hallucinations made significantly more false recognitions than either those suffering from schizophrenia but without verbal hallucinations or healthy controls. There was no difference in the number of false recognitions between healthy controls and those suffering from schizophrenia but not verbal hallucinations (p. 131). This same group of subjects showed a correlation between the frequency of verbal hallucinations and the number of false recognitions on Runs 2 to 4 (Waters, Badcock, Maybery, & Michie, 2003).

Waters and colleagues (2003) also tested this subject group on the Hayling Sentence Completion Test. In this task, subjects are presented with a sentence missing the final word, e.g., “Most cats see well at \_\_\_\_”. The subject’s task is to “complete” the sentence with a word that makes no sense in the context of the sentence, that is, with a word that does not fit grammatically or semantically with the sentence. For example, “goat” for the above example (Burgess & Shallice, 1996). Errors on this test are divided into two types. Type A errors are those that complete the sentence in a grammatically and semantically plausible way, for example “Most cats see well at *night*” (Waters et al., 2003, p. 277). Type B errors are those where the word provided is a possible, but unlikely, end to the sentence, for example “the dough was put in the hot *kitchen*” (p. 277). Waters and colleagues did not perform any by group comparisons on this test; however, they found that

there was a correlation between frequency of verbal hallucinations and the number of Type A errors made by the subject (p. 278). This suggests that the more often a patient hallucinates, the more difficult they find it to inhibit words that most naturally complete the sentences.

Together these results suggest that those suffering from verbal hallucinations have difficulties with the inhibition of thoughts, which is consistent with the model presented here. However, there are other plausible explanations of patients' performance on these tasks, including difficulties with directing attention and automatic inhibition (Paulik et al., 2008). As such, it is not known that abnormal performance on these tasks displayed by those suffering from verbal hallucinations is related to a problem with the intentional inhibition of thoughts and not directed attention or automatic inhibition. However, given previous failures to find such deficits specific to verbal hallucinations (see Kumari et al., 2008, and Perlstein, Carter, Barch, & Baird, 1998, for possible deficits in automatic inhibition related to schizophrenia or schizoaffective disorder but not verbal hallucinations; see Brebion, Smith, Gorman, Malaspina, & Amador, 1998 and Peters et al., 2000, for failures to find deficits with automatic inhibition related to verbal hallucinations), a problem with intentional inhibition is plausibly the cause of these patterns of performance. As such, there is some reason to suppose that the implication of this model that those suffering verbal hallucinations will display deficits with the intentional inhibition of thoughts (such as memories) is borne out.

### ABNORMAL INTENTIONAL INHIBITION WITHOUT A SENSE OF AGENCY DEFICIT

I close by considering a possibly very serious challenge to the model presented here. This challenge comes from the fact that there are two groups of patients who display deficits in intentional inhibition of thought without showing a deficit in their sense of agency over their thoughts. These are patients suffering from orbital-frontal cortex lesions and posttraumatic stress disorder. Those with lesions in the orbital-frontal cortex (OFC) show an increase in false alarms across runs of the inhibition of currently irrelevant memories task (Schnider & Ptak, 1999). These patients do not display deficits in their sense of agency over their thoughts, so it seems that sense of agency deficits do not necessarily arise from deficits in intentional inhibition. Whilst I provide no explanation of these phenomena (that would require a separate mechanism), I suggest that there are good reasons to suppose that these patient's deficits are distinct from those of patients suffering verbal hallucinations.

Badcock and colleagues (2005) suggest that the deficits displayed by the OFC group are not identical to those displayed by the group suffering from verbal hallucinations. They emphasise that those suffering verbal hallucinations do not show an increase in false alarms across runs if the runs are separated by half an hour or more (pp. 130–131). In contrast, those with OFC lesions showed an increase in false alarms across runs even when the runs were separated by an hour (p. 133; Schnider & Ptak, 1999). This difference in performance suggests that there is more than one reason why subjects may show an increase in false alarms across runs on the inhibition of currently irrelevant memories task.

If those suffering from verbal hallucinations and OFC lesions show increases in false alarms across trials for different reasons, then these findings are consistent with the model presented here. On this model there are two (or more) places where deficits would cause problems with the intentional inhibition of thoughts. Only one of these deficits would lead to problems with the sense of agency over one's thoughts: namely, if a subject has problems in the formation of the relevant monitoring representations then that subject should display deficits in both the sense of agency and in intentional inhibition. However, if the subject is able to form the monitoring representations, but is unable to use them to inhibit thoughts then the subject ought to display a deficit in intentional inhibition but not their sense of agency. Keep in mind that on this model it is the *presence* of the monitoring representation used to intentionally inhibit thoughts that elicits the sense of agency: those representations need not actually be deployed (or deployable) to inhibit the thought for the sense of agency to be elicited. It is possible that Schnider and Ptak's (1999) patients' performance on the inhibition of currently irrelevant memories task is due to a deficit in their ability to use (rather than form) monitoring representations. This remains an empirical question and one which could be examined to test the hypothesis that the sense of agency is elicited by these monitoring representations.

Another group of patients who show unusual intentional inhibition of thoughts without a deficit in their sense of agency over their thoughts are those suffering from posttraumatic stress disorder (PTSD). There is evidence that those who suffer from PTSD have difficulty in intentionally inhibiting thoughts and that their attempts to inhibit thoughts do not have long-term effects for emotionally charged thoughts, particularly those related to the trauma that lead to the onset of PTSD. That being said, although worse than controls, those suffering PTSD still have a reasonably robust ability to inhibit unwanted thoughts, but only temporarily. Two studies speak to these conclusions.

Shipherd and Beck (1999) compared the performance of those who lived through sexual assault suffering PTSD and those who lived through the same kind of attack but who weren't suffering from PTSD on a task known

as the “thought suppression paradigm”. For this task, subjects were asked to write down their thoughts over three 9-minute periods following different instructions. In the baseline phase subjects were given no instructions as to what they were to try to think about. The next, called the suppression phase, was like the baseline phase except the patients were told to try not to think about the assault they had suffered. In the final phase, called the expression phase, they were told that they could think of anything including the sexual assault they had suffered (Shipherd & Beck, 1999, pp. 103–104).

Tasks like this have been studied widely. Generally subjects’ performance on this task depends on the emotional salience of what they are told to think about. For things that subjects tend not to think about in the baseline phase and with low emotional salience (e.g., white bears) there is typically a rebound effect such that subjects think about them more following an instruction not to think about them (Merchelbach, Muris, van den Hout, & de Jong, 1991; Wegner, Schneider, Carter, & White, 1987). Once the instruction to inhibit particular thoughts is lifted (for the expression phase) if the thought is emotionally salient then inhibition will continue. However, if the thought is not emotionally salient, then subjects will typically think about it more in the expression phase (Wegner & Gold, 1995).

Comparing the reports of subjects with and without PTSD across each of these three phases provides insight into the intentional inhibition abilities of the two groups. Those suffering PTSD had proportionally more thoughts about the assault than those who did not at baseline (around 30% of reported thoughts as compared to around 15% were assault related). However, during the suppression phase, where subjects were asked to inhibit thoughts about the assault, both groups had the same proportion of assault related thoughts (at around 8%). In the expression phase, where the instruction to inhibited thoughts of the assault was lifted, those suffering from PTSD showed a strong rebound effect with assault related thoughts making up around 20% of their reports. In contrast, those without PTSD continued to inhibit these thoughts in this phase, with roughly the same proportion of thoughts being about the assault as in the suppression phase (Shipherd & Beck, 1999, p. 106).

It seems that efforts to intentionally inhibit emotionally salient thoughts (in this case about a trauma) do not have the same longer term effects for those suffering PTSD as it does for controls. However, it should be noted that a possible confound is that those suffering from PTSD had been attacked more recently than those who were not (Shipherd & Beck, 1999, p. 103)

Cottencin and colleagues (2006) extend these findings by comparing the performance of patients suffering PTSD to healthy controls on a directed forgetting task. In this task, subjects are shown four lists of 24 words in separate runs. Each word is shown for 5 s, followed by an instruction “to

remember” or “to forget” the word (in a proportion of 1:1). Words were not associated with any of the traumas that caused PTSD in any of the patients and were not highly emotionally salient (p. 72). After they were presented with a list of 24 words, subjects were asked to recall as many of the words labelled “to remember” as possible. Thus, they were required to inhibit the words labelled “to forget”. After doing this for all four lists, subjects were given a 10-minute distraction task, then asked to recall as many words as possible regardless of whether they were labelled “to remember” or “to forget” (p. 73).

Those subjects suffering from PTSD recalled fewer words overall during the initial recall phase (in which subjects were asked to recall just the words labelled “to remember”) than healthy controls. Both those suffering from PTSD and healthy controls recalled more words labelled “to remember” than those labelled “to forget”. However, those suffering PTSD recalled more words labelled “to forget” than healthy controls.

Some lessons can be drawn from these results regarding the intentional inhibition abilities of those suffering PTSD. As those suffering PTSD recall more words labelled “to forget” than controls (despite recalling fewer words overall) it seems reasonable to suggest that those suffering PTSD have difficulty inhibiting the words labelled “to forget”. Furthermore, it may be this is a difficulty with intentional inhibition rather than automatic inhibition as those suffering PTSD do not do worse than controls on measures of automatic inhibition, such as the Stroop colour naming task (Cottencin et al., 2006, p. 78). This difficulty in intentional inhibition of thoughts (in this case memories of words) displayed by those suffering PTSD is in marked contrast to the findings of Shipherd and Beck (1999) discussed earlier. However, it is important to note that patients suffering from PTSD are still capable of inhibiting words labelled “to forget” on the directed forgetting task, as they recall far fewer words labelled “to forget” than labelled “to remember”. Given this, it cannot be concluded from the fact that these patients are worse than healthy controls at inhibiting words labelled “to forget” that they have an inability to inhibit unwanted thoughts. However, it may be the case that intentional inhibition is somewhat harder for those suffering PTSD than it is for healthy controls. Taken together, these results seem to suggest that those suffering PTSD are capable of intentionally inhibiting thoughts when required, but (at least when the thoughts are of no emotional significance) these actions may be harder than they are for controls.

In the final recall stage of Cottencin and colleagues’ (2006) study, where subjects were asked to recall as many words as possible regardless of how they were labelled, those suffering from PTSD again recalled fewer words overall than healthy controls. As in the immediate recall phase, both groups

recalled more words labelled “to remember” than “to forget”; however, this difference was smaller for those suffering from PTSD than it was for healthy controls (p. 73). This part of the directed forgetting task is reminiscent of the expression stage of the thought suppression task discussed earlier, in that both require a thought that was once being intentionally inhibited to be expressed. For this final stage controls performed as usual by recalling more words labelled “to forget” than when they were required to inhibit memories of these words in the initial recall phase. Recall that this is typical when the thoughts are emotionally neutral as opposed to emotionally charged (Wegner & Gold, 1995). In this final recall stage those suffering PTSD showed a pattern of performance similar to the healthy controls, by recalling more words labelled “to forget” than they did in the initial recall phase. The difference here is smaller than the difference for the control group (p. 74). However, this smaller difference seems to be largely an effect of the greater number of words labelled “to forget” recalled by those suffering PTSD in the initial recall phase. That is, the closing of the gap seems to be due to these patients recalling more words labelled “to forget”. As such, this can be taken as evidence for a deficit in inhibition.

From these two studies it is reasonable to conclude that those suffering PTSD have some problems with intentional inhibition of thoughts despite not having a deficit in their sense of agency. In particular, the intentional inhibition of unwanted thoughts that have strong emotional significance for the patient does not have the same longer term effects as it does for healthy controls. Additionally, those suffering PTSD find it harder to inhibit thoughts of less emotional salience than healthy controls. Despite these problems, those suffering PTSD are able to temporarily inhibit thoughts of emotional salience (as in the suppression phase of the thought suppression paradigm) and inhibit (albeit to a reduced extent) thoughts of limited emotional salience (as in the initial recall phase of the directed forgetting paradigm). In terms of the model presented here, these last two findings indicate that those suffering PTSD are capable of forming the monitoring representations necessary for the intentional inhibition of thought. As such they can use this representation to elicit a sense of agency. As in the case of those suffering OFC lesions, it is consistent with this model for there to be problems with the intentional inhibition of thoughts without deficits in the sense of agency.

## CONCLUSION

In this paper I have argued for a new metacognitive model of the sense of agency over thoughts. The model I present takes the sense of agency over thoughts to be elicited by monitoring representations of the processes of

thinking that are used in the intentional inhibition of thought. This model provides a powerful explanation of the various losses of the sense of agency in a variety of schizophrenic delusions and hallucinations. It also explains the deficits in the intentional inhibition of thoughts displayed by those with a high hallucination predisposition and it is consistent with the occurrence of inhibition deficits without sense of agency deficits.

Manuscript received 17 May 2010

Revised manuscript received 9 September 2010

First published online 9 December 2011

## REFERENCES

- Badcock, J. C., Waters, F. A. V., Maybery, M. T., & Mitchie, P. T. (2005). Auditory hallucinations: Failure to inhibit irrelevant memories. *Cognitive Neuropsychiatry*, *10*(2), 125–136.
- Bentall, R. P., & Slade, P. D. (1985). Reliability of a scale measuring disposition towards hallucination: A brief report. *Personality and Individual Differences*, *6*, 527–529.
- Brebion, G., Smith, M., Gorman, J., Malaspina, D., & Amador, X. (1998). Resistance to interference and positive symptomology in schizophrenia. *Cognitive Neuropsychiatry*, *3*, 179–190.
- Bortolotti, L., & Broome, M. (2009). A role for ownership and authorship in the analysis of thought insertion. *Phenomenology and the Cognitive Sciences*, *8*, 205–224.
- Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34*(4), 263–273.
- Campbell, J. (1999). Schizophrenia, the space of reasons and thinking as a motor process. *Monist*, *82*(4), 609–625.
- Campbell, J. (2002). The ownership of thoughts. *Philosophy, Psychiatry, Psychology*, *9*(1), 35–39.
- Cangas, A. J., Errasti, J. M., Garcia-Montes, J. M., Alvarez, R., & Ruiz, R. (2006). Metacognitive factors and alterations of attention related to predisposition to hallucinations. *Personality and Individual Differences*, *40*(3), 487–496.
- Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about worry and intrusions: The Metacognitions Questionnaire and its correlates. *Journal of Anxiety Disorders*, *11*(3), 279–296.
- Cottencin, O., Vaiva, G., Huron, C., Devos, P., Ducrocq, F., Jouvent, R., et al. (2006). Directed forgetting in PTSD: A comparative study versus normal controls. *Journal of Psychiatric Research*, *40*, 70–80.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *The American Psychologist*, *34*(10), 906–911.
- Frith, C. D. (1992). *The cognitive neuropsychiatry of schizophrenia*. Hove, UK: Psychology Press.
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14–21.
- Gollwitzer, P. M., & Schaal, B. (1998). Metacognition in action: The importance of implementation intentions. *Personality and Social Psychology Review*, *2*(2), 124–136.
- Graham, G., & Stephens, G. L. (1994). *Mind and mine*. In G. Graham & G. L. Stephens, *Philosophical psychopathology* (pp. 91–110). Cambridge, MA: MIT Press.
- Holmes, E. A., & Steel, C. (2004). Schizotypy: A vulnerability factor for traumatic intrusions. *Journal of Nervous and Mental Disease*, *192*(1), 28–34.
- Jost, J. T., Kruglanski, A. W., & Nelson, T. O. (1998). Social metacognition: An expansionist review. *Personality and Social Psychology Review*, *2*(2), 137–154.

- Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, 9, 149–171.
- Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness* (pp. 289–326). New York, NY: Cambridge University Press.
- Koriat, A., & Levy-Sadot, R. (2000). Conscious and unconscious metacognition: A rejoinder. *Consciousness and Cognition*, 9, 193–202.
- Kumari, V., Peters, E. R., Fannon, D., Premkumar, P., Aasen, I., Cooke, M. A., et al. (2008). Uncontrollable voices and their relationship to gating deficits in schizophrenia. *Schizophrenia Research*, 101, 185–194.
- Lobban, F., Haddock, G., Kinderman, P., & Wells, A. (2002). The role of metacognitive beliefs in auditory hallucinations. *Personality and Individual Differences*, 32, 1351–1363.
- Marzillier, S., & Steel, C. (2007). Positive schizotypy and trauma-related intrusions. *Journal of Nervous and Mental Disease*, 195(1), 60–64.
- Mellor, C. S. (1970). First rank symptoms of schizophrenia. *British Journal of Psychiatry*, 117, 15–23.
- Merchelbach, H., Muris, P., van den Hout, M., & de Jong, P. (1991). Rebound effects of thought suppression: Instruction-dependent? *Behavioural Psychotherapy*, 19, 225–238.
- Nelson, T. O. (1996). Consciousness and metacognition. *The American Psychologist*, 51(2), 102–116.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In: G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125–169). San Diego, CA: Academic Press.
- Paulik, G., Badcock, J. C., & Maybery, M. T. (2006). The multifactorial structure of the predisposition to hallucinate and associations with anxiety, depression and stress. *Personality and Individual Differences*, 41, 1067–1076.
- Paulik, G., Badcock, J. C., & Maybery, M. T. (2007). Poor intentional inhibition in individuals predisposed to hallucinations. *Cognitive Neuropsychiatry*, 12(5), 457–470.
- Paulik, G., Badcock, J. C., & Maybery, M. T. (2008). Dissociating the components of inhibitory control involved in predisposition to hallucinations. *Cognitive Neuropsychiatry*, 13(1), 33–46.
- Perlstein, W. M., Carter, C. S., Barch, D. M., & Baird, J. W. (1998). The Stroop task and attention deficits in schizophrenia: A critical evaluation of card and single-trial Stroop methodologies. *Neuropsychology*, 12, 414–425.
- Peters, E. R., Pickering, A., Kent, A., Glasper, A., Irani, M., David, A., et al. (2000). The relationship between cognitive inhibition and psychotic symptoms. *Journal of Abnormal Psychology*, 109(3), 386–395.
- Proust, J. (2003). Does metacognition necessarily involve metarepresentation? *Behavioural and Brain Sciences*, 26(3), 352.
- Racsmány, M., Conway, M. A., Garab, E. A., Cimner, C., Janka, Z., Kurimay, T., et al. (2008). Disrupted memory inhibition in schizophrenia. *Schizophrenia Research*, 101, 218–224.
- Reder, L. M., & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 45–77). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schneider, A., & Ptak, R. (1999). Spontaneous confabulators fail to suppress currently irrelevant memory traces. *Nature Neuroscience*, 2, 677–681.
- Seth, A. K. (2008). Post-decision wagering measures metacognitive content, not sensory consciousness. *Consciousness and Cognition*, 17(3), 981–983.
- Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition*, 9(2), 313–323.
- Shipherd, J. C., & Beck, J. G. (1999). The effects of suppressing trauma-related thoughts on women with rape-related posttraumatic stress disorder. *Behaviour Research and Therapy*, 37, 99–112.

## 24 CARRUTHERS

- Waters, F. A. V., Badcock, J. C., Maybery, M. T., & Michie, P. T. (2003). Inhibition in schizophrenia: Associations with auditory hallucinations. *Schizophrenia Research*, *62*, 275–280.
- Wegner, D. M., & Gold, D. B. (1995). Fanning old flames: Emotional and cognitive effects of suppressing thoughts of a past relationship. *Journal of Personality and Social Psychology*, *5*, 782–792.
- Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*, 5–13.
- Young, G. (2008). On how a child's awareness of thinking informs explanations of thought insertion. *Consciousness and Cognition*, *17*(3), 848–862.