

Sense of Agency and Higher-Order Cognition: Levels of explanation for schizophrenia¹

Shaun Gallagher

Abstract: I contrast two general approaches to the explanation of certain symptoms of schizophrenia: a top-down model that emphasizes the role of introspection, and a bottom-up model that takes both neurological and phenomenological aspects into account. Top-down accounts look to higher-order cognitive attributions as disrupting a sense of self-agency; bottom-up accounts suggest that the phenomenal sense of agency is generated by specific neuronal mechanisms, and that these mechanisms are disrupted in cases of schizophrenia. The bottom-up view is supported by both clinical reports and experimental science, and is consistent with a phenomenological approach to psychopathology.

Philosophers and cognitive neuroscientists are interested in psychopathologies primarily for theoretical reasons. They often appeal to what goes wrong in pathological cognition to illuminate how things work in "normal" cognition. While psychiatrists are primarily concerned with treatment and therapy, theoreticians can learn from their clinical accounts, and especially from the patient's own narratives about their experience. Phenomenological approaches are characterized as starting with experience rather than with pre-established theories. Phenomenological psychiatrists and philosophers take the patient's first-person narratives seriously. That is, they regard them as reflective of the patient's actual experiences. Experience, itself, however, is complex. One can distinguish between (1) first-order phenomenal experience, that is, the immediately, pre-reflective, lived-through experience of the world, and (2) higher-order cognition, a reflective experience which supports the ability to make attributive judgments about one's own first-order experience. Cognitive neuroscientists are interested in explaining how such experiences are generated in a third level of the cognitive system, (3) the non-conscious, sub-personal processes that are best described as neuronal or brain processes.²

¹ An earlier version of this paper was presented at the European Society for Philosophy and Psychology. Lyon, France (July 2002)

² There may be some intermediate level(s) of description (syntactical or representational) that are also understood as non-conscious, but for purposes of this paper I will leave this

According to what Francisco Varela (1996) has called *neuropsychology*, cognitive neuroscience ought to be guided in some way by the experience that it attempts to explain. A neuroscientist who, for example, is using neuroimaging techniques, will look for very different things if she starts with the idea that schizophrenic symptoms like thought insertion and delusions of control are caused by a dysfunction of higher-order cognition, than if she starts with the idea that schizophrenic symptoms are primarily manifested at the level of first-order experience. These two starting points represent a contrast between a *top-down* explanation and a *bottom-up* explanation of schizophrenic symptoms. In this paper I want to make this distinction clear, and argue that to explain positive symptoms of schizophrenia, like thought insertion and delusions of control, within the phenomenological-neurological framework, one needs to take a bottom-up approach, and that involves attending to the first-person experience of patients.³

Distinctions between self-agency and self-ownership in schizophrenia

Distinctions between self-agency and self-ownership may be found both in first-order phenomenal experience and higher-order consciousness. In regard to the latter, for example, Graham and Stephens (1994) work out their account of introspective alienation in schizophrenic symptoms of thought insertion and delusions of control in terms of two kinds of self-attribution.

- **Attributions of subjectivity (ownership):** the subject reflectively realizes and is able to report that he is moving or thinking. For example, he can say, "This is my body that is moving."
- **Attributions of agency:** the subject reflectively realizes and is able to report that he is the cause of his movement or thinking. For example, he can say "I am causing this action."

possibility aside. Also, I don't deny that there may be unconscious mental states, but I will also leave this idea aside. Assume also that experience is not simply cognitive, but also emotional and embodied.

³Similar distinctions serve as a framework for the analysis found in Gerrans (2001), and they are consistent with his remarks there, and with Gallagher (2000a and 2000b; 2004). The approach taken here fits into the general contours of a phenomenological approach to schizophrenia, the history and main points of which are outlined by Aaron Mishara (National Institute of Mental Health) in an unpublished text, "Agency and Self in Schizophrenia: Should Icarus More Fear the Heights or the Depths?" Mishara distinguishes between an Apollonian (top-down) and Dionysian (bottom-up) approach. For a good example of a clinically informed, phenomenological, bottom-up approach, see Parnas (2003) and Sass (2000, 1998).

This distinction seems consistent with another very similar distinction made at the level of first-order phenomenal consciousness.

- **Sense of ownership:** the pre-reflective experience or sense that I am the subject of the thought or movement ("I am thinking," or "I am moving").
- **Sense of agency:** the pre-reflective experience or sense that I am the cause or author of the thought or movement.

These two distinctions are not identical, even if they are consistent or similar.⁴ They differ in regard to the level or order of self-conscious experience. The distinction between sense of ownership and sense of agency belongs to a first-order phenomenology; it involves a pre-reflective (non-conceptual) sense of ownership and agency implicit to experience, and generated by neurological processes responsible for personal control of action or thought.⁵ In contrast, when one introspectively *attributes* ownership (subjectivity) and agency, one takes a second- or higher-order attitude toward first-order phenomenal experience (see Lambie and Marcel, 2002; Gallagher and Marcel, 1999). Thus, in explanations of schizophrenic symptoms such as delusions of control, thought insertion, and auditory hallucination, these distinctions work at different levels. More generally they help to distinguish between cognitive (higher-order, top-down) versus phenomenological (first-order, bottom-up) accounts of schizophrenia.

The distinctions between ownership and agency at either level can be worked out by considering the difference between voluntary and involuntary movement. If, for example, someone pushes me from behind, I sense that it is my body that is moving (ownership/ subjectivity), but that I did not cause the movement (no agency). This is like the complaint in schizophrenic delusions of control: My body moved (ownership), but I did

⁴The distinction understood in this way is made in Gallagher (2000a, 2000b). I am suggesting that this is not the same distinction found in Graham and Stephens (1994), although in their later work (Stephens and Graham, 2000) they do make a very similar one. As I will make clear in the following, however, Stephens and Graham explain the distinction between a sense of agency and a sense of ownership as a production of higher-order cognition that is read into the level of phenomenal consciousness. Thus, what Stephens and Graham characterize as a sense of agency that is "a normal component or strand in our experience of thinking," but "normally phenomenologically intertwined with introspective awareness as well as with the sense of subjectivity" (2000, p. 9), turns out, on their view, to be "constituted by our self-referential narratives or conceptions of our underlying intentional states" (p. 183). The difference in explanation (and perhaps in phenomenology) between a sense of agency generated "from below" by neurological processes, and a sense of agency generated "from above" by cognitive processes goes to the heart of the matter as it is explicated in this paper.

⁵ Recent studies based on brain imaging suggest the involvement of the right inferior parietal cortex, the anterior insula, and other areas (Chaminade and Decety, 2002; Farrer and Frith 2001) in the generation of the sense of agency. See below.

not cause the movement (no agency); moreover, the subject complains, someone else caused the movement. Frith provides the following example of a patient who attributes his own movement to an alien agency.

'The force moved my lips. I began to speak. The words were made for me'. (Frith, 1992 p. 66).

A similar example is given by Mellor.

'When I reach my hand for the comb it is my hand and arm which move, and my fingers pick up the pen, but I don't control them'. (Mellor 1970, p. 17)

One can also understand thought insertion in these terms. In the ordinary case of involuntary or "unbidden" thoughts (Frankfurt, 1976) I might say that there is a thought in my stream of consciousness, that is, it is a thought that I am experiencing (ownership), but that I did not intend to think this thought (no agency), although in such cases, of course, I can certainly attribute agency to myself. That is, I know at a second-order level that this unbidden thought was generated by me. In contrast, however, the schizophrenic would deny his own agency and may attribute the cause to someone or something else.

'I look out my window and I think that the garden looks nice and the grass looks cool, but the thoughts of Eamonn Andrews come into my mind. There are no thoughts there, only his He treats my mind like a screen and flashes his thoughts onto it like you flash a picture'. (Mellor 1970, p. 17).

In all of these cases, there is a lack of a sense of agency for the action or thought, and for the schizophrenic the problem seems to be about agency rather than about ownership. Schizophrenics who suffer from these symptoms acknowledge that they are the ones that are moving, that the movements are happening to their own body, or that thoughts are happening in their own stream of consciousness, but they claim they are not the agents of these movements or thoughts – when in fact they do cause of the movement or thought.

Problems with top-down accounts

I want to suggest that accounts of problems with self-agency are best developed from the bottom up – specifically, that a neurological account helps to explain a failure manifested in first-order experience, and that just such a failure helps to explain why there is a misattribution of agency at the higher-order level. On this approach, the majority of the

explanation needs to be worked out at the neurological and phenomenological levels, where the latter involves an account of what it's like for the subject – an analysis of the basic first-order phenomenal experience involved in the *senses* of ownership and agency or their disruption. Once these analyses are in place, what happens at the cognitive level of self-attribution (the *attribution* of ownership but not of agency, and the misattribution of agency to another) is straight-forwardly explained as an effect of the lower-order disruptions.

This bottom-up strategy is to be distinguished from a top-down account that bestows causal power in this regard to the higher-order cognitive or attributive level. On the latter approach, the misattribution of agency becomes the *explanans* rather than the *explanandum*. Graham and Stephens (1994; and Stephens and Graham, 2000), for example, attempt to work out this type of top-down account. Following Dennett and Flanagan, they propose an explanation of the sense of agency in terms of “our proclivity for constructing self-referential narratives” which allow us to explain our behavior retrospectively: “such explanations amount to a sort of theory of the person's agency or intentional psychology” (1994, p. 101; Stephens and Graham, 2000, p. 161).

[W]hether I take myself to be the agent of a mental episode depends upon whether I take the occurrence of this episode to be explicable in terms of my underlying intentional states (1994, p. 93).

This top-down account depends on a "theory of mind" approach according to which we reflectively make sense of our actions in terms of our beliefs and desires. So, if a patient does or thinks something for which he has no intentions, beliefs, or desires – mental states that would normally explain or rationalize such actions – the first-order movements or thoughts would not appear as something he intentionally does or thinks. Thus, whether something is to count for me as my action

depends upon whether I take myself to have beliefs and desires of the sort that would rationalize its occurrence in me. If my theory of myself ascribes to me the relevant intentional states, I unproblematically regard this episode as my action. If not, then I must either revise my picture of my intentional states or refuse to acknowledge the episode as my doing. (1994, p. 102).

On this approach, non-schizophrenic first-order phenomenal experience appears the way it does because of properly ordered second-order interpretations, and schizophrenic first-order experience appears the way it does because of second-order *mis-interpretation*.

[T]he subject's sense of agency regarding her thoughts likewise depends on her belief that these mental episodes are expressions of her intentional

states. That is, whether the subject regards an episode of thinking occurring in her psychological history as something she does, as her mental action, depends on whether she finds its occurrence explicable in terms of her theory or story of her own underlying intentional states. (Graham and Stephens, 1994, p. 102; see Stephens and Graham, 2000, pp. 162ff).

It would follow from this view that schizophrenic symptoms are inferential mistakes made on the basis of higher-order introspective or perceptual self-observations. On the account offered by Graham and Stephens, "what is critical is that the subject find her thoughts inexplicable in terms of beliefs about her intentional states" (1994 p. 105). On this theory of mind approach, subjects normally attempt to explain their own experience to themselves. Graham and Stephens suggest that (schizophrenic) mistakes in such explanations may be motivated by negative emotions concerning the thoughts or movements in question, but that they need not be so motivated. They point out that this is complicated by the fact that in many cases, the thoughts/movements are innocuous or neutral. In such cases no evaluative or emotive aspects need be involved. In effect, the failure to attribute agency, or the misattribution of agency may simply be a (mis)judgment made about the incongruence between the content of the current experience and what the subject takes to be her more general conception of her intentional states (see, Stephens and Graham 2000, p. 170).

Graham and Stephens's top-down explanation ignores first-level phenomenology and has nothing to say about neurological processes that may be involved. In line with higher-order-representational theories of consciousness, for Graham and Stephens the first-order experiences of schizophrenic symptoms are not lived through in any originally conscious sense, but are determined by theoretical mistakes made at higher cognitive levels. Thought X seems not to be my thought, only *after* some reflective verification process has failed.

This is not an uncommon type of analysis. Ralph Hoffman (1986) proposes an explanation of verbal hallucinations in schizophrenia that depends on the failure of a self-corrective, introspective judgment that normally corrects the default and automatic inference that ordinary unintended instances of inner speech are caused by something other than myself. On this view, a default mechanism defines non-intentional actions/thoughts as not caused by me – and in schizophrenia this default continues to function. In the non-schizophrenic case, however, on some second-order level, there is a properly functioning "self-corrective process" that vetoes this mechanism and verifies that in fact such actions are my own. One learns that "unintended or alien representations occur during prior passive states and thereby dismiss their veracity" (Hoffman, 1986, p. 509). In schizophrenia this second-order cognitive correction fails and

leads to the mis-attribution.

Stephens and Graham (2000) contest Hoffman's view because it fails to explain why we correctly attribute agency for non-self-generated speech to others when we are in a passive (listening) state and the speech is in fact generated by others. Hoffman's self-corrective process, as he describes it, would prevent us from doing so. But what Stephens and Graham offer as "a more plausible version" is more plausible only if one accepts the idea that there is something like a second-order cognitive deliberation that occurs each time we hear someone else speak or generate our own inner speech. They suggest that the self-corrective process should be thought of as a judgment-withholding process which "induces the subject to withhold judgment on or reconsider such nonself inferences" (p. 107). This second-order cognitive process would have to consider a variety of evidence about whether it is one's own (internal) or someone else's (external) speech.

Although Stephens and Graham do not provide any indications of what may be happening in the brain that would cause the introspective problems in schizophrenia, there are some top-down explanations that are interestingly combined with neurological explanation. For example, Frith's (1992) account of inserted thought appeals to problems on the metarepresentational level, where metarepresentation is defined as a full-fledged second-order act of reflection. The failure of metarepresentational introspection is attributed to neurological dysfunctions associated with sub-personal efferent copy at a brain mechanism called the comparator. It can be argued, however, that Frith's account misconstrues the phenomenological (first-order) level of experience, mistakenly correlates brain mechanisms responsible for first-order experience to second-order cognition, and fails to explain the *mis*-attribution of agency (Gallagher, 2000; Gallagher, 2004b; also see Stephens and Graham, 2000, pp. 141ff).⁶

John Campbell (1999), although claiming to follow Frith, moves, I think, in a slightly different direction. For him, the problem is not with the introspective or metarepresentational level of self-monitoring. Second-order (introspective) processes play no causal role here, and in fact, on his view, second-order self-monitoring functions must continue to work properly even for cases of inserted thought. Campbell, however, like Frith, mixes second-order cognitive processes and neurological processes that correlate with first-order experience,⁷ and yet does not consider the

⁶ Frith now acknowledges the problems in his 1992 account for explaining thought insertion, although not necessarily for delusions of motor control (see Frith 2004).

⁷ Campbell (1999) writes: "It is the match between the thought detected by *introspection*, and the content of the efferent copy picked up by the comparator, that is responsible for the sense of [agency for] the thought. ... You have knowledge of the content of the thought only through introspection. The content of the efferent copy is not itself conscious. But it is match at the monitor [= comparator] between the thought of which

possibility that first-order experience might in fact play a role in the problem.

A bottom-up account

A good example of a phenomenologically-informed bottom-up approach is given by Louis Sass (1992, 1998). His account is quite in contrast to Stephens and Graham.⁸ On Sass's account, neurological problems may cause tacit sensory-motor processes that are normally implicit in first-order phenomenal experience to become abnormally explicit. This becoming explicit is already a form of automatic, or what Sass (2000) calls "operative" hyperreflexivity, and it may motivate more willful forms of hyper-reflective awareness (also see Sass 2003; Sass and Parnas 2003). The normally tacit integration of cognitive, emotional, and motivational experience is disrupted in schizophrenic experience at the phenomenal level; the implicit unity of the self breaks down; and one begins to feel alienated from one's thoughts and actions. For Sass, this more primary disruption often brings on reflective forms of hyperreflexivity involving an excessive type of second-order introspection. Though secondary and defensive in a causal sense, this introspective hyper-reflection often plays an important causal role in bringing on problems of its own. To the extent that higher order processes do play a causal role in bringing about self-alienation, this occurs not as the original source of the problem, but as part of a cascade of processes that begin on a lower level.

Sass thus presents a view of the potentially complex interactions that can occur between more automatic, lower-level processes and higher-level ones that may have a more willful, quasi-willful, and possibly defensive quality. The account that I outline in this section pursues this kind of bottom-up perspective. It takes the distinction between sense of ownership and sense of agency at the level of first-order phenomenal experience seriously, and presents a bottom-up model consistent with recent empirical findings.

On this bottom-up account, problems with self-agency that manifest

you have introspective knowledge and the efferent copy that is responsible for the sense of being the agent of that thought. It is a disturbance in that mechanism that is responsible for the schizophrenic finding that he is introspectively aware of a thought without having the sense of being the agent of that thought."

⁸Stephens and Graham (2000) misinterpret Sass's work in this regard. They view Sass as proposing a top-down explanation, bestowing causal power on hyper-reflective introspection. But Sass (1992) explicitly claims that a higher-order introspection need not be the origin of hyperreflexivity or self-alienation, and indeed, that lower-level processes are probably prior. He attributes a possible causal role to the disruption or under-activation of more automatic and less volitional neurophysiological processes (pp. 69, 386). For Sass's brief critique of Graham and Stephens, see Sass (1999 pp. 261-62).

themselves in thought insertion and delusions of control are generated on a neurological level. The neurological picture is complex, but recent results of brain-imaging studies suggest the importance of two neuronal areas in generating a sense of agency for movement.

Farrer and Frith (2001) have shown contrasting activation in the right inferior parietal cortex for perception of action caused by others, and in the anterior insula bilaterally when action is experienced as caused by oneself. One possible explanation for the involvement of the right inferior parietal cortex in the discrimination between self-agency and other-agency is suggested by Jeannerod (1999). Namely, actions performed by others are perceptually mapped in allocentric coordinates. Farrer and Frith note that "there is strong physiological evidence that the inferior parietal cortex [involves this] kind of remapping ... to generate representations of body movements in allocentric coordinates" (p. 601).

In contrast to the function of the right inferior parietal cortex, Farrer and Frith suggest that the role of the anterior insula in providing a sense of self-agency involves the integration of three kinds of signals generated in self-movement: somatosensory signals (sensory feedback from bodily movement, e.g., proprioception), visual and auditory signals, and corollary discharge associated with motor commands that control movement. "A close correspondence between all these signals helps to give us a sense of agency" (p. 602).⁹

Other studies show that lesions in the right parietal lobe can lead to difficulties in attributing actions. For example, lesions in the right parietal cortex can cause a disturbance of the sense of ownership for one's limbs (as in neglect or alien hand syndrome). Also, in psychiatric and neurological patients self-awareness disorders have been linked to metabolic abnormalities in the right inferior parietal cortex. In schizophrenic patients the feeling of alien control (delusions of control) during a movement task has been associated with an increased activity in the right inferior parietal lobe (Spence, et al., 1997).

Of course things are likely more complicated than this, in both normal and schizophrenic experience. The sense of agency for motor action may depend on a pre-action, forward motor control mechanism that matches motor intention and efference copy of the motor command. The

⁹ Studies by Decety et al (2002), Chaminade and Decety (2002) and Farrer, Franck, Georgieff, Frith, Decety, and Jeannerod (2003) support this conclusion. It is important to note that in the experiments mentioned here the authors have adopted the same concept I have defined above as the sense of agency, and as outlined in Gallagher (2000a). Other empirical studies consistent with the findings mentioned here have also used this definition (see, e.g., Blakemore et al., 2000; Fournieret, et al. 2001; Jeannerod, 2003; Ruby and Decety, 2001; van den Bos and Jeannerod, 2002; Vogeley et al. 2001; Vogeley and Fink, 2003).

proper timing of such a match may depend on the proper functioning of the supplementary motor area, the premotor, and prefrontal cortexes, and such functions are known to be disrupted in schizophrenic subjects with delusions of control (Fournieret and Jeannerod, 1998, Georgieff and Jeannerod, 1998; see Haggard and Eimer, 1999; Haggard and Magno, 1999; Malenka et al., 1982). Moreover, there may be a more general or basic disruption of neuronal processes that affect not just the sense of agency for motor action, but disrupt the sense of agency for cognitive experience (resulting in symptoms of thought insertion). The sense of agency for thought may depend on the anticipatory aspect of working memory, something that may also malfunction in schizophrenic subjects with delusions of control (see, Daprati, et al., 1997; Franck et al., 2001; Singh et al., 1992; Vogeley, et al. 1999).¹⁰

Although it is important to sort out the specific nature of the neurological problem, for purposes of this essay one can say that whatever are the precise nature of the neurological disruptions responsible for delusions of control or thought insertion, some such processes clearly generate a first-order phenomenal experience that lacks a sense of agency. The neurological problems occur in very basic, primary mechanisms for motor control, or working memory, and generate a first-order *sense that I am not the agent* of my movement or thought. And this is just the phenomenology reported by the schizophrenic patient. Clearly, both the supposition of neuroscience and the most reasonable explanations of how experience is generated, suggest that what phenomenal experience is like depends (at least in part) on the proper functioning of the brain. In the case of the schizophrenic who suffers from delusions of control and thought insertion, neurological problems generate a first-order experience that does not include a sense of agency for certain actions and thoughts. The experience, summarized in a first-person, albeit abstract way, is this: "I am not the agent of my movement or thought."

If this in fact is an accurate characterization of schizophrenic experience at the level of first-order phenomenal consciousness, then in

¹⁰ Even if we were reticent to accept the schizophrenic's own reports of such experience, of which there are many, behavioral studies support this interpretation. A simple but elegant example is provided by an experiment conducted by Frith and Done (1988) which demonstrates that schizophrenic patients are unable to distinguish between self-generated or externally generated sounds. A randomly generated and relatively loud tone will elicit a relatively large response in EEG, but if a normal subject generates the tone himself in a self-generated and spontaneous action of pressing a key, the evoked response will be of a smaller magnitude. One-hundred percent of control subjects showed larger responses to externally-generated than to self-generated tones. In contrast, eighty percent of schizophrenic patients were equally startled (as measured by EEG response) in the two conditions. A significant number of patients surprised themselves, and did not experience self-agency in the production of the tone. More recent experiments on the ability of schizophrenic patients to anticipate events show results consistent with this study (see, e.g., Posada et al., 2001).

some respects it seems clear that when a second-order report on such experience is either confused or quite clear about the lack of agency for a particular movement or thought, it is not mistaken. The subject correctly reports her experience, that she experiences herself as the subject (owner) of, but not the agent of the movement; she attributes ownership but not agency for the thought. At higher-order levels of introspective report, or reflective metarepresentation, the subject simply reports what she experiences at the first-order level. This is not a mistake, as suggested by Stephens and Graham, but a report of what the subject actually experiences.

This, however, is not the entire story. If the only problem were a lack of agency, then there would be no difference between the phenomenology of passive, involuntary movement, or unbidden thoughts, and schizophrenic delusions. The explicit delusional component of schizophrenic experience involves the subject's misattribution of agency to another person (or machine, or thing). In this regard, a lack of a sense of self-agency does not add up to an attribution of action to others. Here there appears to be two possible explanations.

- (a) The cause of the misattribution is based on inferences made at the higher-order level of attributive consciousness.

In this case, the subject introspectively misinterprets her experience as caused by someone else. The lack of a sense of agency is filled-in by a productive narrative. Since the subject has no sense of agency for the action, she makes out that it must be someone or something else that is causing her to act that way. The misattribution may be a supplement or a way to deal with the problems found in experience -- a higher-order, personal-level way of coping with what seem to be the facts. It may be generated in an overactive introspection, motivated by the first-order experience of a lack of agency. This is quite consistent with Graham and Stephens' proposal. They suggest that, in regard to inserted thought, the content of the thought may seem to the subject relevant to context, or appropriately intentional, and thus reflective of agency. But since they are not the agent, someone else must be responsible. The alien nature of the thought is thus a conclusion drawn through higher-order considerations. In a similar way, movements of my body that do not seem to belong to me lead me to a conclusion. "I *conclude* that I am possessed, that my movements are directed by the intentional states of another and express his or her beliefs and desires" (Graham and Stephens 1994, pp. 106-107, emphasis added).

Evidence against this top-down account comes from an examination of the effects of abnormal metarepresentation in pathologies other than schizophrenia. Top-down explanations bestow on the kind of metarepresentational introspection found in schizophrenia the power to

generate self-alienation. But if it can be shown that the specific type of higher-order introspective cognition found in schizophrenic patients can also be found in other pathologies that do not involve introspective alienation, then just this kind of introspective cognition would not be sufficient to explain the schizophrenic effects (e.g., the misattribution of agency) manifested in thought insertion and delusions of control. It turns out that one can find pathologies other than schizophrenia (cases of utilization behavior, Anarchic Hand Syndrome [as distinguished from Alien Hand Syndrome (Frith and Gallagher 2002; see reports in Della Sala 2000; Marchetti and Della Sala 1998; Tow and Chua 1998)], and obsessive-compulsive disorder) that manifest similar abnormal forms of metarepresentation, but do not involve introspective alienation (see Gallagher 2004a for discussion).

A second possible explanation can be stated as follows:

- (b) Some neurological component responsible for the differentiation between self and other is disrupted, and as a result, some sense of alterity is already implicit in the first-order experience.

In this case, the attribution of agency to another is not the result of a theoretical stance that would force the subject to infer that since he is not the agent, someone else must be, or a supplemental account generated in introspection, the odd result of a productive narrative; rather, it is a genuine report of what is truly experienced. This is not meant to rule out the fact that odd, paradoxical, and wildly delusional narratives are often generated as the illness develops. The initial motivation for such narratives, however, may be shaped by processes that start out as completely rational at the second-order level – that is, a completely correct report of what the subject experiences.

One can find evidence in support of this second, bottom-up explanation of misattribution in the neuroscience of overlapping neural representations. The brain areas that are activated when I engage in specific intentional action turn out to be in large part the same brain areas that are activated when I observe someone else engage in the same activity.¹¹ A number of researchers suggest that just such overlapping or shared representations may play some part in our ability to simulate the

¹¹ Brain mapping experiments (using PET or fMRI) during action, the imaginary enactment of one's own action, and the observation of another person's action show activation of a partially overlapping cortical and subcortical network that includes structures directly concerned with motor execution (motor cortex, dorsal and ventral premotor cortex, lateral cerebellum, basal ganglia) and areas concerned with action planning (dorsolateral prefrontal cortex and posterior parietal cortex). In the premotor cortex and the supplemental motor area (SMA), the overlap between imagined and observed actions is almost complete (Decety et al, 1997, Grezes and Decety, 2001; Jeannerod, 1999; Jeannerod and Frak, 1999; Jeannerod, 2001).

thoughts and attitudes of others (Blakemore and Decety 2001; Chaminade et al. 2001; Decety 2002; Jeannerod 2001). This suggests that if something goes wrong with these overlapping neural functions, this "Who" system (Georgieff and Jeannerod 1998), our own movement or our own thoughts may be experienced at the first-order phenomenal level as initiated by someone else.¹² There is good evidence that this is what happens in some schizophrenic patients. (Jeannerod, et al., 2003). In such cases, then, not just the lack of a sense of agency, but also the immediate sense of alterity, may be implicit in first-order experience.

Conclusion

Let me conclude with some important qualifications in regard to the bottom-up explanation I've been developing. By defending the idea that neurological disruptions may generate problems with the sense of self-agency at the first-order phenomenal level, and that second-order ascriptions may be correct reports of what the subject actually experiences, I do not want to suggest that I have sketched the entire atiological line that leads to schizophrenic symptoms. Obviously there is more complexity both to the brain and to experience than we can take in here. But more than that, I do not want to rule out the possibility that personal-level phenomena may start the ball rolling, that some emotional or intersubjective event may spark the genetically predisposed brain to shift towards a more schizophrenic dynamics. In many cases, some personal-level aspect seems implicated since the schizophrenic often reports that it is some specific person (or machine or thing) responsible for the movement or thought, and that the inserted thought consists of a specific message or content.¹³ In this

¹² "This relative similarity of neurophysiological mechanisms accounts for both the fact that actions can normally be attributed to their veridical author, and that action attribution remains a fragile process. Indeed, there are in everyday life ambiguous situations where the cues for the sense of agency become degraded and which obviously require a subtle mechanism for signaling the origin of an action" (Jeannerod et al., 2003).

¹³ Beyond the issue of misattribution, there are in fact two unresolved problems concerning thought insertion and delusions of control noted in the clinical literature but rarely addressed in the theoretical literature. One is the *problem of the episodic nature of positive symptoms*: for example, the fact that not all but only some of the schizophrenic's thoughts are experienced as inserted thoughts. That this is the case is clear, not only from empirical reports by patients, but by logical necessity. The subject's complaint that various thoughts are inserted depends on a necessary contrast between thoughts that seem inserted and those that do not seem inserted -- and at a minimum, the thoughts that constitute the subject's introspective complaint cannot seem inserted. If all thoughts were experienced as inserted by others, the subject would not be able to complain "in his own voice," so to speak. The second problem is *the specificity of positive symptoms*. In this regard, in cases of thought insertion, specific kinds of thought contents, but not all kinds appear to be thought inserted. For example, delusional experiences are sometimes associated with specific others. A schizophrenic may report that thoughts are being inserted by a particular person, or that they are always about a specified topic. In auditory hallucination the voice always seems to say the same sort of thing. I argue elsewhere that

regard, aspects of Graham and Stephens's account may play a role in explaining why the subject misattributes agency to a *specific* other -- notably, their suggestion about the effect of emotion may be important. And once underway, hyper-reflective introspection may enhance that shift in dramatic ways, producing the frequently hyperbolic narratives and disrupted thought processes of schizophrenia (see Gallagher 2003). In the end, then, a full explanation is likely to involve a combination of these sub-personal and personal factors.

References

- Campbell, J. 1999. Schizophrenia, the space of reasons and thinking as a motor process. *The Monist*, 82 (4): 609-625.
- Chaminade, T. and Decety, J. 2002. Leader or follower? Involvement of the inferior parietal lobule in agency. *Neuroreport* 13 (1528): 1975-78.
- Danckert, J., Ferber, S., Doherty T., Steinmetz, H. Nicolle, D. and Goodale, M. A. 2002. Selective, non-lateralized impairment of motor imagery following right parietal damage. *Neurocase* 8: 194-204.
- Decety J, Chaminade T, Grèzes, J. and Meltzoff, A.N. 2002. A PET Exploration of the Neural Mechanisms Involved in Reciprocal Imitation. *Neuroimage* 15: 265-272.
- Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. and Fazio, F. 1997. Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain* 120: 1763-77.
- Della Sala, S. 2000. Anarchic hand: the syndrome of disowned actions. *Creating Sparks, The BA Festival of Science*. www.creatingsparks.co.uk.
- Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Daléry, J., Jeannerod, M. 1997. Looking for the agent: An investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*, 65, 71-96.
- Farrer, C. and Frith, C.D. 2001. Experiencing oneself vs. another person as being the cause of an action: the neural correlates of the experience of agency. *NeuroImage* 15: 596-603.
- Fish, F. J. 1985. *Clinical Psychopathology: Signs and Symptoms in Psychiatry*. Ed. M. Hamilton. Wright.
- Fourneret, P. and Jeannerod, M. 1998. Limited conscious monitoring of motor performance in normal subjects, *Neuropsychologia* 36: 1133-1140.
- Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Daléry, J., d'Amato, T., and Jeannerod, M. 2001. Defective recognition of one's

explanations that remain totally on the sub-personal level will not be able to address these two problems (Gallagher 2004b).

- own actions in patients with schizophrenia. *American Journal of Psychiatry*, 158: 454-59.
- Frankfurt, H. 1976. Identification and externality. In A. O. Rorty (ed). *The Identities of Persons* (pp. 239-51). Berkeley: University of California Press.
- Frith, C. D. 2004. Comments on Shaun Gallagher. *Psychopathology*, 37 (2004): 20-22.
- Frith, C. D. 1992. *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Frith, C. D., and Done, D. J. 1988. Towards a neuropsychology of schizophrenia. *British Journal of Psychiatry*, 153: 437- 443.
- Frith, C. and Gallagher, S. 2002. Models of the pathological mind. *Journal of Consciousness Studies*, 9 (4): 57-80.
- Gallagher, S. 2004a. Agency, ownership and alien control in schizophrenia. In P. Bovet, J. Parnas, and D. Zahavi (eds). *The structure and development of self-consciousness: Interdisciplinary perspectives* (pp. 89-104) Amsterdam: John Benjamins Publishers.
- Gallagher, S. 2004b. Neurocognitive models of schizophrenia: A neurophenomenological critique. *Psychopathology* 37: 8-19.
- Gallagher, S. 2003. Self-narrative in schizophrenia. In A. S. David and T., Kircher (eds.). *The Self in Neuroscience and Psychiatry* (pp. 336-357). Cambridge: Cambridge University Press.
- Gallagher, S. 2000a. Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences* 4 (1): 14-21
- Gallagher, S. 2000b. Self-reference and schizophrenia: A cognitive model of immunity to error through misidentification. In D. Zahavi (ed). *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-experience* (pp. 203-239). Amsterdam & Philadelphia: John Benjamins.
- Gallagher, S. and A. J. Marcel. 1999. The Self in Contextualized Action. *Journal of Consciousness Studies* 6 (4): 4-30.
- Garrens, P. 2001. Authorship and ownership of thoughts. *Philosophy, Psychiatry, & Psychology* 8.2/3: 231-237
- Georgieff, N. and Jeannerod, M. 1998. Beyond consciousness of external events: A Who system for consciousness of action and self-consciousness. *Consciousness and Cognition*, 7, pp. 465–77.
- Graham, G. and Stephens, G. L. 1994. Mind and mine. In G. Graham and G. L. Stephens (eds). *Philosophical Psychopathology* (pp. 91-109). Cambridge, MA: MIT Press.
- Grèzes, J. and Decety, J. 2001. Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, 12, pp. 1–19.
- Haggard, P. and Eimer, M. 1999. On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research* 126: 128-33.

- Haggard, P. and Magno, E. 1999. Localising awareness of action with transcranial magnetic stimulation. *Experimental Brain Research* 127: 102-107.
- Hoffman, R. 1986. Verbal hallucinations and language production processes in schizophrenia. *Behavioral and Brain Sciences* 9: 503-517.
- Jeannerod, M. 1999. To act or not to act: perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology, A* 52:1-29
- Jeannerod, M. & Frak, V. 1999. Mental imaging of motor activity in humans. *Current Opinions in Neurobiology*, 9:735-9
- Jeannerod, M. 2001. Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, 14, S103-S109.
- Jeannerod, M., Farrer, C., Franck, N., Fournier, P., Posada, A., Daprati, E. Georgieff, N. 2003. Action recognition in normal and schizophrenic subjects. In: T. Kircher and A. David (eds). *The Self in Schizophrenia: A Neuropsychological Perspective*. Cambridge: Cambridge University Press.
- Junginger, J. 1986. Distinctiveness, unintendedness, location, and non-self attribution of verbal hallucinations. *Behavioral and Brain Sciences* 9: 527-28.
- Lambie, J. A. and Marcel, A. J. 2002. Consciousness and the varieties of emotion experience: A theoretical framework. *Psychological Review* 109 (2): 219-259.
- Malenka, R. C., Angel, R. W., Hampton, B. and Berger, P. A. 1982. Impaired central error correcting behaviour in schizophrenia. *Archives of General Psychiatry*, 39: 101-107.
- Marchetti, C. and Della Sala, S. 1998. Disentangling the Alien and Anarchic Hand. *Cognitive Neuropsychiatry* 3(3): 191 - 207
- Mellor, C. S. 1970. First rank symptoms of schizophrenia. *British J Psychiatr*, 117: 15-23.
- Parnas, J. 2003. Anomalous self-experience in early schizophrenia: A clinical perspective In: T. Kircher and A. David (eds). *The Self in Schizophrenia: A Neuropsychological Perspective*, Cambridge: Cambridge University Press.
- Posada, A., Franck, N., Georgieff, N. & Jeannerod, M. (2001) Anticipating incoming events: An impaired cognitive process in schizophrenia. *Cognition*, 81, 209-225.
- Sass, L. 2003. Schizophrenia and the self: hyper-reflexivity and diminished self-affectation. In: T. Kircher and A. David (eds). *The Self in Schizophrenia: A Neuropsychological Perspective*, Cambridge: Cambridge University Press.
- Sass, L. 2000. Schizophrenia, self-experience, and the so-called negative symptoms. In D. Zahavi (ed). *Exploring the Self* (pp. 149-82). Amsterdam: John Benjamins.
- Sass, L. 1999. Analyzing and deconstructing psychopathology. *Theory and Psychology* 9 (2): 257-268.

- Sass, L. 1998. Schizophrenia, self-consciousness and the modern mind. *Journal of Consciousness Studies*, 5: 543-65.
- Sass, L. 1992. *Madness and Modernism: Insanity in the Light of Modern Art, Literature, and Thought*. New York: Basic Books.
- Sass, L. and Parnas, J. (2003-in press). Schizophrenia, consciousness, and the self. *Schizophrenia Bulletin*.
- Singh, J. R., Knight, T., Rosenlicht, N., Kotun, J. M., Beckley, D. J. and Woods, D. L. 1992. Abnormal premovement brain potentials in schizophrenia. *Schizophrenia Research*, 8: 31-41.
- Spence SA, Brooks DJ, Hirsch SR Liddle, PF Meehan, J and Grasby PM. 1997. A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control) *Brain* 120: 1997–2011.
- Stephens, G. L. and Graham, G. 2000. *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts*. Cambridge, MA: MIT Press.
- Varela, F. J. 1996. Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, 3 (4): 330-49.
- Vogeley, K., Kurthen, M., Falkai, P., Maier, W. 1999. The human self construct and prefrontal cortex in schizophrenia. The Association for the Scientific Study of Consciousness: Electronic Seminar (<http://www.phil.vt.edu/assc/esem.html>).
- Tow, A. M. and Chua, H. C. 1998. The alien hand sign--case report and review of the literature. *Ann-Acad-Med-Singapore*, 27(4): 582-5