**Somatoparaphrenia, Anosognosia, and Higher-Order Thoughts**

Rocco J. Gennaro

[Final version in *Disturbed Consciousness: New Essays on Psychopathology and Theories of Consciousness*, ed. Rocco J. Gennaro, The MIT Press, 2015.]

Somatoparaphrenia is a pathology of self characterized by the sense of alienaton from parts of one’s body. It is usually construed as a kind of delusional disorder caused by extensive right hemisphere lesions. Lesions in the temporoparietal junction are common in somatoparaphrenia but deep cortical regions (for example, the posterior insula) and subcortical regions (for example, the basal ganglia) are also sometimes implicated (Valler and Ronschi 2009). Patients are often described as feeling that a limb belongs to another person and thus attribute ownership of the limb and bodily sensation to someone else. There is also some question as to whether or not the higher-order thought (HOT) theory of consciousness can plausibly account for the depersonalization psychopathology of somatoparaphrenia (Liang and Lane 2009, Rosenthal 2010, Lane and Liang 2010). Liang and Lane argue that it cannot. The HOT theory of consciousness says that what makes a mental state a conscious mental state is that it is the target of a HOT to the effect that “I am in mental state M” (Rosenthal 2005, Gennaro 2012). When the HOT is itself is unconscious, the conscious state is still outer-directed. When the HOT is conscious, we have *introspection* and so the conscious thought is directed at the mental state. In section I, I briefly review the previous exchange between Lane and Liang and David Rosenthal. In section II, I further explore somatoparaphrenia and the nature of delusion while offering a number of additional replies to Lane and Liang. In section III, I examine the central notions of “mental state ownership” and “self-concepts” in an effort to account especially for the depersonalization aspect of somatoparaphrenia against the background of HOT theory. In section IV, I argue that to the extent that somatoparaphrenia casts doubt on the notion that some thoughts are “immune to error through misidentification” (IEM), the most fundamental aspect of IEM is still consistent with HOT theory. Overall, I argue that HOT theory is left unscathed by the pheneomenon of somatoparaphrenia and can even help to explain what happens in these cases.

**I. Lane and Liang vs. Rosenthal: an overview**

Liang and Lane (2009) initially argued that somatoparaphrenia threatens HOT theory because it contradicts the notion that according to HOT theory, when I am in a conscious state, I have the accompanying HOT that “I am in mental state M.” The ‘I’ is not only importantly *self-referential* but essential in tying the conscious state to *oneself* and thus to one’s *ownership* of M. Indeed, it is difficult to understand how one can have a conscious state but not, at least implicitly, attribute it to oneself.

Rosenthal (2010) basically responds that one can be aware of bodily sensations in two ways which, normally at least, go together: (a) aware of a bodily sensation *as one’s own*, and (b) aware of a bodily sensation *as having some bodily location*, like a hand or foot. Patients with somatoparaphrenia still experience the sensation as their own but also as having a mistaken bodily location (perhaps somewhat analogous to phantom limb pain). Such patients still do have the awareness in (a), which is the main issue at hand, but they have the very strange awareness in sense (b). So somatoparaphrenia leads some people to misidentify the bodily location of a sensation as someone else’s but the awareness of the sensation itself remains one’s own.

Rosenthal also accepts what he calls a “more modest” version of Shoemaker’s (1968) immunity to error through misidentification (IEM) principle which says that if the ground of my judgment is introspective, that is “from the inside,” whenever I say or think for example that “I feel pain” it cannot be the case that I am mistaken in thinking that the person in pain is me. The same goes for other mental states and bodily sensations. According to Rosenthal, we should at least adopt a “thin immunity principle” (TIP) such that when I have a conscious pain (for example) I cannot be wrong about whether it’s I who I think is in pain (Rosenthal 2005, 357). Elsewhere, he explains that “no error is possible about whom I am aware of as having the pain because the spontaneous awareness tacitly identifies the bearer of the pain with the bearer of the awareness” (Rosenthal 2010, 274). I will return to this theme later in section IV.

But Lane and Liang (2010) are not satisfied and, among other things, counter that Rosnethal’s analogy to phantom limbs is faulty and that he has still not explained why the identification of the *bearer* of the pain cannot also go astray especially since Rosenthal clearly holds that misrepresentation can occur between a HOT and its object.

**II. The Nature of Delusion and Further Replies to Lane and Liang**

Although I largely agree with much of Rosenthal’s response, I believe that there are many responses available to a HOT theorist that have thus far been neglected in this debate. I will focus on these further replies in this section with special attention to nature of delusion.

*First*, we must remember that many of these patients often deny feeling *anything* in the limb in question (Bottini *et al*. 2002). As Liang and Lane point out, patient FB, while blindfolded, feels “no tactile sensation” (2009, 664) when the examiner would in fact touch the dorsal surface of FB’s hand. In these cases, it is therefore difficult to see what the problem is for HOT theory at all. While somatoparaphrenia is indeed a puzzling phenomenon, the HOT theorist can simply reply that there is no HOT in such cases because there is no conscious feeling (and vice versa). Since there is no conscious feeling, there is no allegedly problematic “I” in a HOT that potentially conflicts what the patient says or feels. Moreover, HOT theory can *explain why* there is no conscious feeling, that is, because there is no HOT (with its “I-concept”) directed at a mental state.

*Second*, in those cases where FB does report feeling a tactile sensation, she is told that the examiner is about to touch her *niece’s hand*. When FB is required to report touches on her niece’s hand, FB’s tactile perception increased dramatically but she insisted that she was feeling touches on someone else’s hand. Of course, if FB is not really feeling anything, then the reply above would hold.

However, in cases where there really is a bodily sensation of some kind, a HOT theorist might plausibly argue that there are really *two* conscious states which seem to be at odds. There is a conscious feeling in a limb but also the (conscious) attribution of the limb to someone else. However, it is also crucial to emphasize that somatoparaphrenia is often characterized as a *delusion* of *belief*, often under the broader category of *anosognosia*, a condition in which a person who suffers from a disability seems unaware of the existence of the disability (Breen *et al*. 2000, de Vignemont 2010, Feinberg 2011). A delusion is often defined as a false belief held based an incorrect (and probably unconscious) *inference* about external reality or oneself that is firmly sustained despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary (Radden 2011, Bortolotti 2013). In some cases, delusions seriously inhibit normal day-to-day functioning. This doxastic conception of delusion is common among psychologists and psychiatrists (Bayne and Pacherie 2005, Bortolotti 2009). Beliefs, generally speaking, are themselves often taken to be intentional states integrated with other beliefs. They are typically understood as caused by perceptions or experiences which then lead to action or behavior. Thus, somatoparaphrenia is, in some ways, closer to self-deception and involves frequent confabulation. If this is a reasonable interpretation of the data, then a HOT theorist can argue that, in these cases, the patient has the following *two* conscious states:

(1) S1: A conscious *feeling* (i.e. a tactile sensation) in the limb in question, and

(2) S2: A conscious *belief* that the limb (and thus sensation) belongs to someone else.

Now having both S1 and S2 simultaneously might be strange and perhaps even self-contradictory in some sense, but the puzzlement has nothing to do with HOT theory. Indeed, this possibility is perfectly consistent with HOT theory. S1 would still require an unmediated HOT that “I am in M.” There is no problem here for HOT theory because it is still true that one is having a conscious sensation in these cases. Note that the HOT itself is *unconscious* in these cases so that the patient is not consciously entertaining the HOT or “I” in question. But what is really odd is that the person with somatoparaphrenia *also* has, as S2 indicates, a conscious belief which runs counter to or even contradicts S1. FB responds to questions by affirming this belief. But surely no theory of consciousness can automatically rule out such a bizarre combination of conscious states, and there is nothing in HOT theory which causes any special difficulty here. HOT theory is a theory about what makes mental states conscious and is not itself a theory of mental content, let alone a theory about the consistency of mental content. If we demand that a theory of consciousness close off the possibility of a pair of contradictory conscious states or an inconsistent pair of a feeling and verbal report, then it is unclear how any theory could pass this test. In short, HOT theory can be true even when a subject exhibits irrational thinking or behavior.

*Third*, it is of course normally true that when a person P feels something in one’s body, P will also believe that the limb is P’s, that is, one’s own limb. Normally, there is presumably an unconscious inference or, perhaps even better, the general *presupposition* that the limb is one’s own. For HOT theory, this presupposition is precisely embodied in the unmediated accompanying unconscious HOT. In cases of somatoparaphrenia, however, there might *also* be an erroneous or confabulated *inference* from the feeling in question to the belief that the sensation and limb are not one’s own. Indeed, according to Feinberg’s (2010, 2011) “ego disequilibrium theory,” delusional misidentification syndromes such as somatoparaphrenia, in part, result from pathological defenses including projection. To be sure, this is abnormal and difficult to understand to some extent, but it is not a problem specifically for HOT theory as far as I can see. These patients *do* consciously feel something which can still be explained by HOT theory in the usual way. But, in addition, there is at least initially an erroneous inference to the delusional belief that the sensation is not one’s own. This belief is evident in FB’s very strange and clearly confabulatory verbal responses to questions. So what would the inference be like? Based on the above discussion, it would presumably take something like the following form:

(1) I am having (or feel) a bodily sensation

(2) I do not have the limb where I feel this sensation

(3) *Therefore*, I feel a sensation in someone else’s limb.

Further confabulation may explain who I choose as the bearer of this sensation. In the case of FB, she chooses her niece based on the examiner’s questions. Yet Liang and Lane think that somatoparaphrenia presents a particular difficulty for HOT theory. But, interestingly, in Lane and Liang’s counter-reply to Rosenthal, they actually say that “FB came *to believe* that her left hand belonged to her niece” (Lane and Liang’s 2010, 497, added emphasis).

So I agree with Rosenthal that there are two different attributions, namely, ownership and bodily location. I agree that “FB is aware of the sensation as being her own. But she is also *aware of* the sensation as having a subjective bodily location in a hand that is not part of her own body” (Rosenthal 2010, 272, emphasis added). For the reasons given above, however, it seems better to construe the ‘aware of’ in the previous sentence as ‘believes that.’ Lane and Liang (2010) are probably right the Rosenthal relies too heavily on the phantom limb analogy since in phantom limb who feels the pain is not at issue. Nonetheless, I think that Lane and Liang do not properly recognize that the tactile sensation at issue is still *felt by FB* as her own in one sense even though she *also* attributes her limb to her niece, and thus believes that the tactile sensation is her niece’s.

Liang and Lane tell us that “what seems to be happening is that these tactile sensations are *represented as belonging to someone other than self*” (2009, 664) and that is problematic for HOT theory. But this is highly ambiguous because another way to “represent” sensations as belonging to someone else is via a propositional attitude such as a belief. And there would be no problem for HOT theory as such as to whether or not these patients can have the conscious belief in S2. That is, a patient with somatoparaphrenia would still represent that *belief* as her own. And the fact remains that no mental state of *her niece’s* is made conscious by any HOT of FB’s. It is certainly not the case that *FB’s niece* has a conscious sensation due to one of FB’s HOTs. Further, as Billon and Kriegel point out (this volume), there are no cases of patients with somatoparaphrenia claiming to *feel a sensation* that is not theirs. These last points still hold despite blindfolded FB’s somewhat odd results on so-called “catch trials” in the Bottini *et al*. study (2002, 251). I am not quite sure that we should take all of these results at face value especially given what I have argued above, but FB does report touches when she was cued to expect her niece’s hand would be touched and when she was actually touched. However, FB reported nothing when she was cued to expect her niece’s hand to be touched and her hand was not touched. Bottini et al (2002, 251) explain that on each trial, “the examiner briefly touched the dorsal surface of FB’s hand” and when she was cued to expect her niece’s hand to be touched, “the examiner touched FB’s left hand.”

It might be objected at this point that I have ignored the “endorsement” and “explanationist” theories of delusion formation or perhaps have assumed the latter in the laying out the inference above. The endorsement account holds that delusions are formed by endorsing as veridical the content of the unusual experience (Bayne and Pacherie 2004) whereas the explanationist account holds that delusions are formed as a way to explain an unusual experience (Maher 1999). Patients with anosognosia and somatoparaphrenia are also arguably well motivated to conjure up an explanation, for example, as a way to manage negative emotions (see McKay and Kinsbourne 2010 for much more on this line of argument). One way to frame the above two theories of delusion formation is by asking the following question: Does the delusional belief occur before the person has the conscious experience or afterward? Alternatively: are delusions bizarre convictions that alter one's way of seeing the world, or are they hypotheses formulated to account for some unusual experiences, and then endorsed as beliefs?

But, as Langdon and Bayne (2010) point out, most delusions are likely hybrids of both views especially when considered over periods of time. They instead propose a “continuum” from “received” to “reflective” delusion whereby the former largely come to pass via the endorsement process and the latter are formed mainly by an explanationist processes. Indeed, it seems to me that the explanationist account makes far more sense when the subject is first asked to respond to questions, such as in FB’s initial examinations. Langdon and Bayne (2010) also note that we should distinguish between “spontaneous” and “provoked” confabulation whereby the latter arises only in response to direct questioning (Kopelman 2010). After all, what makes many patient dialogues so compelling and bizarre is the clear confabulations of the patient in response to questions. For example, when FB was asked how she could report touches on someone else’s hand, her “response was initially elusive; however, she eventually explained that [her] absent-minded niece would always forget her hand on [FB’s] bed while leaving the hospital…” (Bottini *et al* 2002, 251). However, once the delusional belief has taken firm hold in the patient, it then seems more plausible from that point onward to suppose that the delusional belief (and the inference) in question is already in place prior to subsequent experiences, which is more in line with the endorsement theory. The notion that cognitive states can have an impact on one’s very experiences is fairly widely acknowledged.[[1]](#endnote-1)

So if the argument above is plausible with regard to an initial provoked explanationist account, then the subsequent endorsement is more likely to occur unconsciously. This is because the more conceptually loaded resulting experience reflected in (3) above is the *result* or *consequence* of having a delusional belief and the requisite lack of self-awareness.

It is worth mentioning that something like the above analysis has been proposed as a parallel way to account for Capgras syndrome. Bortolotti (2013) explains:

For some, it is correct to say that the delusional belief *explains* the experience. Others claim that the delusion is an *endorsement* of the experience. According to the *explanationist* account (Maher 1999), the content of experience is vaguer than the content of the delusion, and the delusion plays the role of one potential explanation for the experience. For instance, in the Capgras delusion, the experience would be that of someone looking very much like my sister but not being my sister. The delusion would be an explanation of the fact that the woman looks like my sister, but her face feels strange to me: the woman must be an impostor….According to the rival account, the *endorsement* account (Bayne and Pacherie 2004), the content of the experience is already as conceptually rich as the content of the delusion. The delusion is not an explanation of the experience, but an endorsement of it: the content of the experience is taken as veridical and believed. In Capgras, the experience is that of a woman looking very much like my sister but being an impostor, and when the experience is endorsed, it becomes the delusional belief that my sister has been replaced by an impostor.[[2]](#endnote-2)

*Fourth*, to follow up on the points above, it seems to me that when one has a conscious belief, such as S2, we primarily have in mind *introspecting* a belief, that is, consciously thinking about a belief. For example, FB is reporting this (false) belief in response to questions. If this is plausible, then at least some of sting may be taken out of the idea that one can have S2 while also having S1 with its accompanying *unconscious* thought that “I am feeling a sensation.” This is because S1 and S2 may not always be consciously present at the same time in these patients. That is, one may not be consciously aware of both S1 and S2 *at the same time*. Further, the HOT required for S1 is not conscious at all while the belief in S2 is conscious. So it is unclear that whatever inconsistency that might exist in FB’s mind would be as obvious *to her* as we might think it is.

Indeed, to the extent that somatoparaphrenia is typically thought of as a form of anosognosia, it seems plausible to suppose that it too involves a deficit of self-awareness and thus the unity of consciousness in some sense (including perhaps some memory problems as well). Recall that anosognosia involves the lack of awareness of an impairment, such as paralysis to the left side of their bodies. Bayne points out that anosognosic patients “are often oblivious to major changes in the contents of their own conscious states” and appear “to be impaired in tracking [their] own states of consciousness [and] a breakdown in the unity of *reflexive* consciousness” (Bayne 2011, 153, cf. Nikolinakos 2004). This bolsters the above line of argument in the sense that it can explain why FB and others with somatoparaphrenia seem to hold inconsistent beliefs or respond to questions in a way that would seem inconsistent with their beliefs (such as with S1 and S2). That is, they may not be able to hold S1 and S2 reflectively in mind at the same time due to a deficit in self-awareness or introspection. Nikolinakos (2004, 316) explains that “Reflexive consciousness refers to the awareness of phenomenal experience [which] … is a second-order consciousness about information that appears in non-reflexive consciousness,” so it is clear that by ‘reflexive consciousness’ he is referring to what I have been calling ‘introspection.’ This lack of monitoring or tracking of one’s mental states, according to HOT theory, could be explained by an inability to form *conscious* HOTs directed at mental states, that is, as a deficit of introspective ability. What makes anosognosia so puzzling is that we would expect a subject to give up one of two inconsistent beliefs when they are clearly pointed out to that subject, but yet they do not do so. Perhaps the reason is that, from the point of view of the patient, there is an inability to hold in mind introspectively both beliefs at the same time. Much the same might be true in the case of Anton’s syndrome which is a form of anosognosia in which a person with partial or total blindness denies being visually impaired, despite clear medical evidence to the contrary. Just as one with the somatoparaphrenia confabulates in response to questions about the presence of body parts, the Anton’s syndrome patient typically confabulates excuses for the inability to see.

In any case, when Liang and Lane say that “patients typically *feel* that a contralesional limb belongs to someone other than self” (Liang and Lane 2009, 664, emphasis added), this, to my ear, sounds more like a belief rather than some kind of bodily sensation, especially during initial questioning and subsequent confabulation. At the very least, ‘feel’ is ambiguous between a conscious feeling as such and the conscious belief S2. Once again, I suggest that it is the latter partly because FB is asked to verbally respond to questions. The term ‘feel’ in the above quotation seems like the generic all encompassing ‘feel’ such as when one says “I feel that capital punishment is wrong” or “I feel that I can do a better job than Joe.” In these cases, the term ‘feel’ is used more like ‘believe’ or ‘think.’ Indeed, even just the “feel *that*” locution seems to indicate a propositional attitude rather than a bodily sensation like when one says “I feel pain” or “I feel cold” or “I feel tired.” And when Rosenthal states that “being conscious of a state as belonging to someone other than oneself would plainly not make it a conscious state” (as quoted in Liang and Lane 2009, 662), I take it that he is again referring to a possible case where person A believes or ‘feels’ that M is in person B. But, in such a case, M would not become conscious in person B, according to HOT theory. Even if FB does have a HOT about someone *else’s* mental state M based on an inference, M would not become conscious. Thus, the above line of reply, which admittedly reinterprets some of FB’s reports, can be successful *contra* Liang and Lane’s claim (2009, 666).

We must also remind ourselves that HOT theory comes with a well-known “noninferentiality condition” such that a HOT must become aware of its target mental state noninferentially, that is, in an unmediated way. As Rosenthal repeatedly emphasizes, the point of this condition is mainly to rule out certain alleged counterexamples to HOT theory, such as cases where I become aware of my unconscious desire to kill my boss because I have consciously inferred it from a session with my psychiatrist, or where my envy becomes conscious after making inferences based on my own behavior. The characteristic feel of such a conscious desire or envy may be absent in these cases but, since awareness of them arose via conscious inference, the HOT theorist accounts for them by adding this noninferentiality condition. Thus, HOT theory requires that the HOT arises in an unmediated manner. This is also important because if there is any kind of inference to a belief, such as in S2, then the HOT would not arise in the requisite manner. So when Rosenthal says for example that “being aware of oneself as being in pain consists in being aware, in a spontaneous, seemingly unmediated way, of an individual’s being in pain…” (2010, 274), he is referring to an essential and important aspect of standard HOT theory. Lane and Liang are incorrect is treating this appeal to “spontaneous, unmediated awareness” as begging the question with regard to mental ownership (2010, 498). It is an essential aspect of HOT theory. In a more recent paper, Rosenthal explains that “When you are aware of me as being in pain, it will not be the seemingly unmediated awareness characteristic of one’s awareness of one’s own pain, but we are nonetheless aware of the very same thing as being in pain” (2012, 40).

*Fifth*, another potential problem for Lane and Liang can also be seen by recognizing that a HOT is supposed to be a thought *about a mental state*, not about one’s own (or another’s) body or body part. According to HOT theory, when I am in a conscious state, I have the accompanying HOT that “I am in *mental state M*.” Of course, just how bodily sensations should be categorized is a difficult issue in its own right (de Vignemont 2011, section 3), but it at least seems prima facie plausible to construe any of FB’s thoughts about her limbs or her niece’s limbs to be, at least in part, about a perceptual object akin to a perception. In short, we must distinguish between one’s awareness of a mental state or sensation “from the inside,” such as a pain or other tactile sensation, and one’s awareness of a body part, such as a limb. It is true that one need not visually perceive one’s own limb to be aware of it in various ways, such as its spatial orientation via proprioception or where one is having pain. However, Lane and Liang do not acknowledge the fact that *if* some of FB’s thoughts are not really even directed at mental states, then much of what they say is irrelevant with respect to any deep problem for HOT theory. We would no longer have an alleged problem with regard to an “I-thought” *directed at a mental state*, but instead would merely have a first-order conscious thought or belief about a body part or another’s body part (more on this below in section IV).

**III. I-Concepts and Mental State Ownership**

Recall that somoatoparaphrenia is a depersonalization psychopathology as is anosognosia (Prigatano 2010). But what exactly is meant by “depersonalization”? At minimum, it involves a distortion in one’s self-awareness or sense of self. So there is a deficient sense of oneself and distorted self-concepts (Sierra 2009). This seems to include deficits of mental state “ownership.” Let us look more closely at these aspects of somatoparaphrenia.[[3]](#endnote-3)

It is useful to distinguish the *feeling* of ownership from the *judgment* of ownership (de Vignemont 2007, 2011) which is neglected in Lane and Liang’s arguments. As is well-known in cases of phantom pain, one may feel a pain in a limb that does not even exist. On the other hand, in cases of anesthesia or body integrity identity disorder, one might not feel any sensation in a limb which is clearly judged to be part of one’s own body. So mental state “ownership” need not correlate with one’s feeling of “embodiment” (de Vignemont 2007, 2011). In the now well-known rubber hand illusion, one can arguable even feel an external object (a fake rubber hand) as one’s own in some sense. People can be convinced that a rubber hand is their own by putting it on a table in front of them while stroking it in the same way as their real hands.

Along similar lines, Bermudez (2011, pp. 161-166) distinguishes between a “sense” of ownership and a “judgment” of ownership. Recall also that a HOT theorist does not hold that HOTs are typically themselves conscious. Bermudez argues, in the end, that the “sense” of ownership is best viewed as a “judgment” or “thought” rather than as a feeling itself. In contrast to de Vignemont (2007), Bermudez rejects the “inflationary” conception of the sense of ownership according to which there is a distinctive positive phenomenology of ownership, as opposed to the “deflationary” conception which consists in facts about the phenomenology of the bodily sensations and in ownership judgments about the body. “There are facts about the phenomenology of bodily awareness…and there are judgments of ownership, but there is no additional feeling of ownership” (Bermudez 2011, 166). There are still of course bodily sensations and proprioceptive states. What one actually feels is the *first-order* conscious state accompanied by the unconscious judgment or thought about the mental state, as opposed to any phenomenology of “myness.” Notice that this fits nicely with HOT theory which can *explain why* there is a phenomenological sense of myness when one *introspects*, namely, that the HOT is itself conscious whereas no such sense is present when one has an unconscious HOT. The concept “I” is part of a *conscious* thought in the introspective case but part of an *unconscious* thought in the first-order case. Nonetheless, it is certainly true that when there is a disturbance or abnormality in one’s I-concept, such as one’s bodily representation, one’s consciousness will be altered and result in some very odd beliefs and feelings of body *dis*ownership. But, like Bermudez, I do not find it compelling to argue that if a deficit of bodily awareness is manifested in consciousness, then that aspect of bodily awareness is always or even normally part of our consciousness, albeit even in some peripheral way. Many abnormalities of bodily awareness can surely negatively *affect* one’s consciousness but in a way where the corresponding normal functioning would not typically be part of one’s consciousness. “There is no particular reason for understanding a feeling of disownership as the absence of a feeling of ownership – at least, not without prior reasons for thinking that there is such a thing as the feeling of ownership” (Bermudez 2011, 163).

I have used something like this line of response against Ford and Smith’s (2006) argument in favor of the so-called “self-representationalist” theory of consciousness whereby first-order conscious states are always accompanied by an inner-directed peripheral conscious awareness (Kriegel 2009). Ford and Smith contend that cases of depersonalization show that something like Kriegel’s view is correct based on such abnormal cases. But again just because the removal of something—for example, normal proprioception—causes deficits in one’s conscious mental states, it surely does not follow that the awareness of that thing is part of normal conscious experience. The relation could be causal instead of constitutive. That is, the typical abilities and awareness in question might merely, in the normal case, causally contribute to the phenomenology of one’s conscious mental states without being part of the conscious state itself, even peripherally. There are many ways that normal consciousness can be disturbed or impaired (e.g. being unable to breathe) but surely we shouldn’t conclude that every such disturbance shows that the ability in question normally shows up in our phenomenology (see Gennaro 2012, chapter five, especially 127-129).

Let us pursue this general theme a bit further. In previous publications, I have distinguished four degrees of “I-concepts,” the simplest of which is something like “I *qua* this body as opposed to other things” (Gennaro 1993, 1996). Other more sophisticated self-concepts include “I *qua* experiencer of mental states”, “I *qua* enduring thinking thing,” and “I *qua* thinker among other thinkers.” So bodily self-awareness is one of many ways to represent oneself. But it is clear that even this most basic I-concept is ambiguous between, for example, what has been called the “body schema” which has to do with representations implicit in motor control and action and the “body image” which is one’s conscious perceptions of and beliefs about one’s body (Gallagher 1986). There is more than one way to represent one’s body which is presumably realized in distinct neural pathways although the matter is not so simple (de Vignemont 2010, 2011). It is also possible, for example, to have anosognosia *without* somatoparaphrenia. These anosognosics seem to have an intact body image but a distorted body schema in the sense that their denial (or unawareness) of paralysis wrongly leads them to suppose that they are capable of normal motor action. Patients with somatoparaphrenia *also* seem to have deficit in body schema for similar reasons. They too have a severely deficient body image in that they even deny that one of their arms or legs is theirs in the first place (and thus believe that it is someone else’s). For example, FB believes that she can perform certain motor tasks even though she is physically unable to do so. In other cases, a patient might insist that she can clap (or has just clapped) despite left arm paralysis and no ability to move her hand (e.g. patient CC discussed in Berti *et al*. 1998).[[4]](#endnote-4)

Ramachandran (1996) also reports the case of a woman (FD) who suffered from a right hemisphere stroke resulting in left hemiplegia. FD could not move her left arm. But when she was asked to engage in activities which require both hands, such as clapping, she claimed that she could. Ramachandran advances the hypothesis that behaviors giving rise to confabulations and delusions are an exaggeration of normal defense mechanisms that have an adaptive and protective function (Hirstein 2005). So one reason to think that delusions and self-deception overlap is that at least some delusions, like extreme cases of self-deception, appear to have a protective and adaptive function (for much more on this theme, see the essays in Bayne and Fernàndez 2008). Another view about the overlap of delusions and self-deception is that the very existence of delusions supports the traditional account of self-deception such that a person has two contradictory beliefs but is only aware of one of them because she is motivated to remain unaware of the other (McKay *et al*. 2005, 314). This account was also suggested by some of our discussion near the end of the previous section. Levy (2008), for example, argues that the case of FD shows that a person can, at the same time, believe that her arm is paralyzed and believe that she can move her arm. But Levy rightly suggests that this ‘awareness’ comes in degrees. Many people with paralysis and anosognosia simply find ways to avoid tasks which require mobility and then confabulate and make excuses for their lack of behavior in certain contexts.

**IV. Immunity to Error through Misidentification (IEM)**

Let us return to the aforementioned and much discussed immunity to error through misidentification (IEM) principle. According to Shoemaker, a certain subset of thoughts about oneself is immune to error through misidentification (see also Evans 1982). As Shoemaker makes clear, one can think about oneself under any number of descriptions. But only some “I-thoughts” are immune to error through misidentification, namely, those I-thoughts that are directed at one’s mind and mental life, as opposed to one’s body and corporeal life. Wittgenstein (1958/1969) observed that I can see in the mirror a tangle of arms and mistakenly take the nicest one to be mine. I may think to myself “I have a nice arm.” In that case, I may not only be wrong about whether my arm is nice, but also about whom it is that has a nice arm. Such an I-thought about my body (or body part) is not immune to error through misidentification (Kriegel 2007). In extreme abnormal cases, such as mirror self-misidentification, one might even believe that *one's own* reflection in a mirror is some other person.

Recall that even when Lane and Liang say that FB was feeling touches on someone else’s hand, it is still the case that *she* is feeling something. So, in at least this narrow sense, there is still some level of self-awareness such that a self-referential “I” is required for the HOT which accompanies that conscious feeling. The attribution of that feeling or limb to someone else is, as I have urged, best understood as a separate false belief. But notice that there is still no possibility of error with regard to who is having the tactile sensation itself and who is aware of the feeling. Recall Rosenthal’s “thin immunity principle” (TIP) which simply says that when I have a conscious pain (for example) I cannot be wrong about whether it’s I who I think is in pain (Rosenthal 2005, 357). “No error is possible about whom I am aware of as having the pain because the spontaneous awareness tacitly identifies the bearer of the pain with the bearer of the awareness” (Rosenthal 2010, 274). TIP thus holds for FB in the sense that when FB does have a conscious feeling, FB cannot be wrong about whether it is FB who is *aware of* that feeling. The unmediated HOT in question tacitly identifies the bearer of the feeling (FB) with the bearer of the awareness of the feeling (FB). FB cannot be wrong about whether it is FB who FB thinks is having that feeling. So “the identification of the bearer of the pain cannot also go astray” because whatever else FB says or does with respect to the feeling, FB still (unconsciously or implicitly) identifies herself as having that feeling. We might say that FB still “owns” that feeling “from the inside” but she erroneously also attributes it to another person partly due to the delusional belief that the limb is not hers. So when FB insists that she is feeling touches on someone else’s hand, the fact remains that this *feeling itself* is still taken by FB to be her own. Unlike, say, mirror self-identification which involves outward perception directed at one’s body, one cannot be mistaken about the mere awareness of one’s own sensation strictly from the inside. FB actually doesn’t really say that the feeling is not hers but rather that she is feeling touches in someone else’s limb.

Another way to approach this matter is via the somewhat different notion of “bodily immunity to error,” i.e. that certain judgments subjects make about their own bodies based on information gained “from the inside” or “first-person point of view” exhibit IEM (de Vignemont 2012). The notion of “bodily self-ascription” is construed as ascription of bodily properties. De Vignemont (2012) favors an “inside mode account” such that bodily self-ascriptions are immune to error through misidentification if and only if one gains information about the body from the inside (e.g. proprioception, sense of pressure, and sense of balance). Perhaps I can be mistaken that my arms are in fact crossed but it is difficult to see how I could be wrong that the arms that I feel are *mine*. In somatoparaphrenia, there is clearly a failure of bodily ascription since the patient will deny ownership of the limb and attribute the limb to someone else. As we have seen, the delusional aspect of somatoparaphrenia is so strong that patients will confabulate when confronted with clear counter-evidence. However, the counter-evidence is presented via a visual modality and not “from the inside,” such as pointing to the patient’s arm or hand and asking whose it is. When we have some kind of feeling, sensation, or pain, we have cases of ascribing psychological properties, not bodily properties.

One might also distinguish between two ways of violating the IEM principle. Depersonalization and somatoparaphrenia do not result in *false positive errors*, i.e. errors of identification in which “one self-ascribes properties that are instantiated in another individual’s body” (de Vignemont, 2012, 229). Instead, they are at best *false negatives*, i.e. errors of identification in which “one does not self-ascribe properties that are instantiated in one’s own body” (de Vignemont, 2012, 229), though we have seen that even this type of error does not necessarily cause trouble for TIP because there is *also* an awareness of a feeling in one’s own body. So although IEM is normally thought of in terms of the self-ascription of psychological properties, some have explored the possibility that some physical self-ascriptions also have IEM (Evans 1982), such as kinesthetic sensations or proprioception, and even Shoemaker discussed what he called “circumstantial” (as opposed to “absolute”) immunity with respect to physical self-ascriptions depending upon what grounds the ascription is made. But there is a fine line here. On the one hand, bodily *sensations* are still mental states, such as pains and feelings in a limb, and, on the other hand, there are also judgments or feelings about the location and movement of parts of my body (Chen 2009). However, when we discover cases where a patient really *feels*, from the first-person point of view, that, say, one’s left arm is moving when it is not really moving, there seems to be that much more room to doubt that some form of IEM even applies here.[[5]](#endnote-5)

Finally, let’s revisit the notion of “misrepresentation” and HOTs. Recall that Lane and Liang protest that Rosenthal has still not explained why the identification of the bearer of the pain cannot also go astray especially since Rosenthal clearly holds that misrepresentation can occur between a HOT and its target. But whatever one thinks of standard cases of misrepresentation between the first-order and second-order level on HOT theory, they are not clearly relevant here because those cases involve differences in the *contents* of the two respective states.[[6]](#endnote-6) Although Lane and Liang claim (2010, 499) that there should equally be the possibility of a mismatch between the “I” in the HOT and the “I” in the first-order mental state, it is unclear to me how this could be so. Wittgenstein himself usefully distinguished between the “I-as-subject” (e.g. “I have a pain”) and the “I-as-object” (“I have a broken arm”). There is never an “I”-as-object in the *content* of the *first-order* state but there is an implicit (and unconscious) “I”-as-subject at the second-order level *as well as* an “I”-as-object in a typical HOT. According to HOT theory, there would only be an “I-as-subject” concept in any first-order state and the content of the state refers to the outer world. This would be a kind of “raw bearer” of the state, as Rosenthal calls it. After all, if we assume that any mental state must have a bearer, then even first-order states should involve some primitive concept of “I”. The same is true for the unconscious HOT which accompanies a first-order conscious state, but here there is *also* an “I-as-object” referenced in the *content* of the HOT (i.e. “I think that *I am in M*”). Still, these I-concepts are normally parts of unconscious thoughts and so there is little reason to suppose that there is any phenomenological sense of “myness” in these cases. However, when one introspects and has a conscious HOT directed at a mental state, there is not only a conscious “I-as-subject” concept but also a *conscious* “I-as-object” concept in in the content of the HOT which can account for any subjective sense of “myness.” The fact remains, though, that there could be no mismatch between an “I-as-object” in the contents of a mental state M and its HOT because there isn’t even an “I-as-object” concept in the content of M.[[7]](#endnote-7)

I conclude that HOT theory can withstand the alleged threat from cases of somatoparaphrenia. Indeed, I think that HOT theory can even help to explain what happens in these cases, especially when one is clear about the nature of delusions and is careful about the concepts in question.[[8]](#endnote-8)

**References**:

Bayne, T. 2011. *The Unity of Consciousness*. New York: Oxford University Press.

Bayne, T. and Fernández, J. eds. 2009. *Delusions and Self-deception: Affective and Motivational Influences on Belief-Formation*. Hove: Psychology Press.

Bayne, T. and Pacherie, E. 2004. Bottom up or top down? *Philosophy, Psychiatry, & Psychology* 11: 1–11.

Bayne, T. and Pacherie, E. 2005. In defence of the doxastic conception of delusion. *Mind and Language* 20: 163–188.

Bermudez, J. 2011. Bodily awareness and self-consciousness. In S. Gallagher ed. *The Oxford Handbook of the Self*. New York: Oxford University Press.

Berti, A., Ladavas, E., Stracciari, A., Giannarelli, C., and Ossola, A. 1998. Anosognoisa for motor impairment and dissociations with patients’ evaluation of the disorder: theoretical considerations. *Cognitive Neuropsychiatry* 3: 21-44.

Billon, A. and Kriegel, U. (this volume) Jaspers’ dilemma: The psychopathological challenge to subjectivity theories of consciousness.

Bortolotti, L. 2009. *Delusions and Other Irrational Beliefs*. Oxford: Oxford University Press.

Bortolotti, Lisa, “Delusion”, *The Stanford Encyclopedia of Philosophy (Winter 2013 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2013/entries/delusion/>.

Bottini, G., Bisiach, E., Sterzi, R., and Vallar, G. 2002. Feeling touches in someone else’s hand. *NeuroReport* 13: 249-252.

Breen, N., Caine, D., Coltheart, M., Hendy, J. and Roberts, C. 2000. Towards an understanding of delusions of misidentification: four case studies. *Mind and Language* 15: 74-110.

Carruthers, G. 2009. Is the body schema sufficient for the sense of embodiment? An alternative to de Vignemont’s model. *Philosophical Psychology* 22: 123-142.

Chen, C. 2009. Bodily awareness and immunity to error through misidentification. *European Journal of Philosophy* 19: 21-38.

Coltheart, M. 2005. Conscious experience and delusional belief. *Philosophy, Psychiatry & Psychology* 12: 153–157.

Davies, M., Davies, A., and Coltheart, M. 2005. Anosognosia and the two-factor theory of delusions. *Mind and Language* 20: 209-236.

de Vignemont, F. 2007. Habeas Corpus: The sense of ownership of one’s own body. *Mind and Language* 22: 427-449.

de Vignemont, F. 2010. Body schema and body image – Pros and cons. *Neuropsychologia* 48: 669-680.

de Vignemont, Frédérique, "Bodily Awareness", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/bodily-awareness/>.

de Vignemont, F. 2012. Bodily immunity to error. In S. Prosser and F. Recanati eds. *Immunity to Error through Misidentification: New Essays*. Cambridge University Press.

Evans, G. 1982. *Varieties of Reference*. Oxford: Oxford University Press.

Feinberg, T. and Keenan, J. eds. 2005. *The Lost Self: Pathologies of the Brain and Identity*. New York: Oxford University Press.

Feinberg, T. 2010. Neuropathologies of the self: A general theory. *Neuropsychoanalysis* 12: 133-158.

Feinberg, T. 2011. Neuropathologies of the self: clinical and anatomincal features. *Consciousness and Cognition* 20: 75-81.

Ford, J. and Smith, D.W. 2006. Consciousness, self, and attention. In U. Kriegel and K. Williford eds. *Self-Representational Approaches to Consciousness*. Cambridge, MA: MIT Press.

Gallagher, S. 1986. Body image and body schema: a conceptual clarification. *Journal of Mind and Behavior* 7: 541-554.

Gennaro, R. 1993. Brute experience and the higher-order thought theory of consciousness. *Philosophical Papers* 22: 51 – 69.

Gennaro, R. 1996. *Consciousness and Self-consciousness*. Amsterdam and Philadelphia: John Benjamins Publishers.

Gennaro, R. 2006. Between pure self-referentialism and the (extrinsic) HOT theory of consciousness. In U. Kriegel and K. Williford eds. *Self-Representational Approaches to Consciousness*. Cambridge, MA: MIT Press.

Gennaro, R. 2012. *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thhoughts*. Cambridge, MA: The MIT Press.

Hirstein, W. 2005. *Brain Fiction: Self-deception and the Riddle of Confabulation*. Cambridge, MA: MIT Press.

Hirstein, W. ed. 2009. *Confabulation: Views from Neuroscience, Psychiatry, Psychology, and Philosophy*. Oxford: Oxford University Press.

Kopelman, M. 2010. Varieties of confabulation and delusion. *Cognitive Neuropsychiatry* 15: 14-37.

Kriegel, U. 2007. Self-Consciousness. *The Internet Encyclopedia of Philosophy*, Available at: <http://www.iep.utm.edu/self-con/#SH5a>.

Kriegel, U. 2009. *Subjective Consciousness*. New York: Oxford University Press.

Lane, T. and Liang, C. 2008. Higher-order thought and the problem of radical confabulation. *The Southern Journal of Philosophy* 46: 69-98.

Lane, T. and Liang, C. 2010. Mental ownership and higher-order thought. *Analysis* 70: 496-501.

Lane, T. and Liang, C. 2011. Self-consciousness and immunity. *Journal of Philosophy* 108: 78-99.

Langdon, R. and T. Bayne. 2010. Delusion and confabulation: Mistakes of perceiving, remembering and believing. *Cognitive Neuropsychology* 15: 319-345.

Levy, N. 2009. Self-deception without thought experiments. In T. Bayne and J. Fernández eds. *Delusions and Self-deception: Affective and Motivational Influences on Belief-Formation*. Hove: Psychology Press.

Liang, C. and Lane, T. 2009. Higher-order thought and pathological self: the case of somatoparaphrenia. *Analysis* 69: 661-668.

Maher, B.A. 1999. Anomalous experience in everyday life: Its significance for psychopathology. *The Monist* 82: 547–70.

McKay, R., Langdon, R., and Colheart, M. 2005. Sleights of mind: Delusions, defences, and self-deception. *Cognitive Neuropsychology* 10: 305-326.

McKay, R. and Kinsbourne, M. 2010. Confabulation, delusion, and anosognosia: Motivational factors and false claims. *Cognitive Neuropsychology* 15: 288-318.

Nikolinakos, D. 2004. Anosognosia and the unity of consciousness. *Philosophical Studies* 119: 315-342.

Prigatano, G. ed. 2010. *The Study of Anosognosia*. New York: Oxford University Press.

Prosser, S. and Recanati, F. eds. 2012. *Immunity to Error through Misidentification: New Essays*. Cambridge University Press.

Radden, J. 2011. *On Delusion*. New York: Routledge.

Ramachandran, V.S. 1996. The evolutionary biology of self-deception, laughter, dreaming and depression: some clues from anosognosia. *Medical Hypotheses* 47: 347–362.

Rosenthal, D. 2005. *Consciousness and Mind*. New York: Oxford University Press.

Rosenthal, D. 2010. Consciousness, the self and bodily location. *Analysis* 70: 270-276.

Rosenthal, D. 2012. Awareness and identification of self. In J. Liu and J. Perry eds. *Consciousness and the Self: New Essays*. Cambridge, MA: Cambridge University Press.

Shoemaker, S. 1968. Self-reference and self-awareness. *Journal of Philosophy* 65: 555-567.

Sierra, M. 2009. *Depersonalization: A New Look at a Neglected Syndrome*. New York: Cambridge University Press.

Vallar, G. and Ronchi, R. 2009. Somatoparaphrenia: a body delusion. A review of the neuropsychological literature. *Experimental Brain Research* 192: 533-551.

Wittgenstein, L. 1958/1969. *The Blue and Brown Books, 2nd edition*. Oxford: Basil Backwell.

1. For an excellent anthology on confabulation, see Hirstein 2009. [↑](#endnote-ref-1)
2. There is also the much discussed two-factor theory of confabulation which I think I can simply take for granted in this discussion. Coltheart explains the two main factors involved in the formation of delusions as follows: “There is a first neuropsychological impairment that presents the patient with new (and false data), and the delusional belief formed is one which, if true, would explain these data. The nature of this impairment varies from patient to patient. There is a second neuropsychological impairment, of a belief evaluation system, which prevents the patient from rejecting the newly formed belief even though there is much evidence against it. This impairment is the same in all people with monothematic delusions” (Coltheart 2005, 154). I won’t elaborate on this here, but see Davies *et al*. 2005 for some discussion of anosognosia and the two-factor theory of delusions. [↑](#endnote-ref-2)
3. For an excellent anthology of readings exploring numerous “self-related” pathologies, see Feinberg and Keenan 2005. [↑](#endnote-ref-3)
4. See also Carruthers 2009 for a nice critical discussion of de Vignemont’s view and for more on the body image vs. body schema distinction. [↑](#endnote-ref-4)
5. For much more on all things IEM, see the essays in Prosser and Recanati 2012. [↑](#endnote-ref-5)
6. I disagree with the way that Rosenthal handles these cases but won’t pursue this here. See Gennaro 2012, especially chapter four. Lane and Liang (2008) also take Rosenthal to task on so-called “radical confabulation” or “targetless HOT” cases. Although we may disagree about what a HOT theorist should say about possible cases of mismatches or misrepresentation between the HOT and the first-order states, we must be very careful here. Although there is most certainly no infallibility in *introspection*, there is a much more intimate connection between an unconscious HOT and its first-order target state. [↑](#endnote-ref-6)
7. For more on IEM and somatoparaphrenia, see Lane and Liang 2011, especially 84-90, but I think the points made here apply equally to that discussion. [↑](#endnote-ref-7)
8. Thanks to Tim Lane for helpful comments on and a conversation about an earlier version of this paper. [↑](#endnote-ref-8)