

The Same-Order Monitoring Theory of Consciousness

Uriah Kriegel, University of Arizona

In U. Kriegel and K.W. Williford (eds.), *Self-Representational Approaches to Consciousness* (pp. 143-170), Cambridge MA: MIT Press, 2006

1. Introduction

One of the promising approaches to the problem of consciousness has been the Higher-Order Monitoring Theory of Consciousness. According to the Higher-Order Monitoring Theory, what it is like to be in a conscious state, and there being anything it is like to be in it, are determined by the way the state is represented to the subject. More precisely, a mental state *M* of a subject *S* is conscious iff *S* has another mental state, *M**, such that *M** is an appropriate representation of *M* (Armstrong 1968, 1981; Lycan 1996, 2001; Mellor 1978; Rosenthal 1986, 1990, 2002, forthcoming).

Typically, the conscious state and its representation are construed as logically independent of each other: *M* could occur in the absence of *M**, and vice versa. Recently, however, several philosophers have developed a Higher-Order Monitoring theory with a twist. The twist is that *M* and *M** entertain some kind of *constitutive relation*, or *internal relation*, or some other *non-contingent* relation, so that they are *not* logically independent of each other. For want of a better term, I will call this the *Same-Order Monitoring Theory of Consciousness*.¹ For the sake of clarity, I will reserve the name “Higher-Order Monitoring” to the standard version that insists on the logical independence of *M* and *M**.

Versions of the Same-Order Monitoring Theory can be found in recent writings by Brook and Raymont (Forthcoming Ch. 5), Carruthers (2000 Ch. 9),

Caston (2002), Gennaro (1996 Ch. 2, 2002, 2004b), Hossack (2002, 2003), Kobes (1995), Kriegel (2002a, 2003a, 2003b, 2005), Lehrer (1996a, 1996b, 1997 Ch. 7, 2002, 2004), Lurz (2003a, 2003b, 2004), Natsoulas (1993a, 1996a, 1996b, 1999, 2004), Sanford (1984), Smith (1986, 1989 Ch. 2, 2004), Thomasson (2000), Van Gulick (2001, 2004), and Williford (2003b, 2005). Despite this surprising number of proponents, there has not been among philosophers of mind an explicit awareness of the emerging appeal of this new and distinctive approach to consciousness – at least not *as such*, i.e., *as* a new and distinctive approach.

In this paper, I will first expound and then propound the Same-Order Monitoring Theory (SOMT). The paper accordingly divides in two. §§2-3 attempt primarily to articulate the basic idea behind SOMT and to offer an exposition of several accounts of consciousness from the recent philosophical literature that are effectively versions of SOMT, as well as to assess the relative plausibility of these different versions. §§4-5 argue the superiority of SOMT over the Higher-Order Monitoring Theory (HOMT), by developing two major difficulties for HOMT that do not apply to SOMT.²

Naturally, the concerns of the present paper will be of interest primarily to those who have at least *some* trust in the monitoring approach to consciousness and who find at least *some* merit in standard HOMT. But I hope that the discussion of the subtler developments of the approach will interest also those with no sympathy for it, if only because doing away with the monitoring approach to consciousness would presumably require squaring off with its best version. It is therefore worthwhile to consider what the best version is.

2. The Same-Order Monitoring Theory

HOMT and SOMT agree that the presence of a higher-order representation of M is a *necessary* condition for M's being conscious. The standard argument for this goes something like this (see Lycan 2001):³

1. Mental states the subject is completely unaware of are unconscious states; so,
2. If mental state M of a subject S is conscious, then S must be aware of M; but,
3. Awareness of something involves a representation of it; therefore,
4. If M is conscious, then S must have a representation of M.

It is clear, however, that the presence of a higher-order representation is not a *sufficient* condition for M's being conscious.⁴ This is why the monitoring approach appeals to the notion of an "appropriate" or "suitable" higher-order representation: even though the presence of a higher-order representation is not a sufficient condition for M's being conscious, the presence of an *appropriate* higher-order representation is. The question is what makes a given higher-order representation "appropriate" in the relevant sense.

This is where versions of the monitoring approach differ. There are several dimensions along which these various versions contrast with each other. Perhaps the most widely acknowledged distinction is between versions that construe the higher-order representation as perception-like and versions that construe it as thought-like. Thus, according to Rosenthal, a higher-order representation is appropriate only if it is a thought, whereas according to Armstrong and Lycan, it must be a quasi-perceptual state. This distinction leads to a contrast between Higher-Order Thought (HOT) Theory and Higher-Order Perception (HOP) Theory.

The contrast that concerns us here is between HOMT and SOMT. According to SOMT, one of the requirements on an “appropriate” higher-order representation is that it bear some constitutive relation, some logical connection, to M; HOMT rejects this requirement.⁵ By way of introducing such a requirement, Kobes (1995: 294) writes:

[Let us introduce] a token constituency requirement: the first-order mental state or event must, at the time that it is conscious, be a constituent part of the HOT [higher-order thought] event token.

In similar vein, Van Gulick (2001: 295) writes:

Although both HOP and HOT theorists assume distinctness or nonidentity [of the monitored state and the monitoring state]... one could try to develop the higher-order view in a way that rejected or at least weakened that assumption...

Let us make this contrast explicit by formulating the two competing accounts as follows:

(HOMT) A mental state M of a subject S is conscious iff S has a mental state M*, such that (i) M* is an appropriate representation of M, and (ii) there is *no* constitutive relation between M and M*.

(SOMT) A mental state M of a subject S is conscious iff S has a mental state M*, such that (i) M* is an appropriate representation of M, and (ii) there *is* a constitutive relation between M and M*.

That is, SOMT postulates an internal, non-contingent relation between S's conscious state and her awareness of her conscious state. HOMT construes these two as completely logically independent.

Different constitutive relations define different versions of SOMT.⁶ The strongest constitutive relation is of course *identity*. Accordingly, the strongest version of SOMT holds that M is identical with its higher-order representation. This means, in effect, that M represents *itself*.⁷ The view can be formulated as follows:

(SOMT₁) A mental state M of a subject S is conscious iff S has a mental state M*, such that (i) M* is an appropriate representation of M, and (ii) M = M*.

Which is equivalent to:

(SOMT₁') A mental state M of a subject S is conscious iff M is a representation of itself.⁸

This sort of view has been recently defended by Smith (1986, 1989 Ch. 2, 2004),⁹ Thomasson (2000), Caston (2002), Kriegel (2003b), Hossack (2002, 2003), Williford (2003), Brook (Ms), Raymont (Ms), and Brook and Raymont (forthcoming).¹⁰

Another version of SOMT appeals to the *part-whole* relation, also a constitutive relation. (That it is a constitutive relation can be seen from the fact that some philosophers – e.g., Armstrong (1978), Lewis (1991) – conceive of the part-whole relation as *partial identity*.) On this view, for a mental state to be conscious, it is not sufficient that the subject be aware of it; the subject's awareness of it must be *part of* that very same mental state. A view of this sort is

defended by Gennaro (1996 Ch. 2, 2002, 2004b), Van Gulick (2001, 2004), and Kriegel (2002a, 2003b, 2005). It may be formulated as follows:

(SOMT₂) A mental state M of a subject S is conscious iff S has a mental state M*, such that (i) M* is an appropriate representation of M, and (ii) M* is a (proper) *part of* M.

The relevant notion of parthood here is not spatial or temporal, but *logical*.¹¹ There are complicated questions surrounding the explication of the notion of logical parthood, questions to which justice cannot be done here.¹² But an example may suffice to illustrate the nature of the logical-part whole relation. When I am *glad* that the weather is nice, I necessarily also *believe* that the weather is nice; it is impossible to be glad that the weather is nice without believing that this is so. But my belief that the weather is nice is not an extra mental act, which occurs *in addition to* my gladness. Rather, the belief is somehow *inherent in*, or *built into*, the gladness. In other words, my belief is *part of* my gladness, in a logical sense of “part of.” So my believing that the weather is nice is a *logical part* of my being glad that the weather is nice.¹³ Likewise, according to SOMT₂, our awareness of our conscious states is a *logical part* of those conscious states. When I have a conscious experience of blue, I am aware of my conscious experience. But the awareness is not an extra mental act, which occurs *in addition to* the experience. Rather, the awareness is *inherent in* – it is *built into* – the experience. It is in this sense, then, that M* is claimed in SOMT₂ to be a logical (proper) part of M.

In the above formulation of SOMT₂, it is explicitly required that M* be a *proper* part of M. This is intended to insure that SOMT₂ be exclusive of SOMT₁.¹⁴ But it is significant that a version of SOMT can be formulated that will

remain silent on whether M^* is a proper or improper part of M , thus covering both $SOMT_1$ and $SOMT_2$. This version would be formulated as follows:

($SOMT_3$) A mental state M of a subject S is conscious iff S has a mental state M^* , such that (i) M^* is an appropriate representation of M , and (ii) M^* is a (*proper or improper*) part of M .

When M^* is a *proper* part of M , $SOMT_3$ accords with $SOMT_2$; when it is an *improper* part, $SOMT_3$ accords with $SOMT_1$. But $SOMT_3$ allows both structures to be involved in conscious states.

A close neighbor of $SOMT_2$ is the idea that for M to be conscious is for it to have *two* parts, such that one represents the other. The view may be formulated as follows:

($SOMT_4$) A mental state M of a subject S is conscious iff (i) M^* is a (*proper*) part of M , (ii) M^\diamond is a (*proper*) part of M , and (iii) M^* is an appropriate representation of M^\diamond .

The idea here is, in a way, that the conscious state involves a “mereological sum” of the monitoring state and the monitored state. (Again, the relevant notion of mereology is that of *logical* mereology, not spatial or temporal mereology.¹⁵) This is to be distinguished from $HOMT$, in which the conscious state is identified with the monitored state solely.¹⁶

Again, $SOMT_4$ is formulated in term of proper parthood in order to make it exclude the previous versions of $SOMT$, but more general versions can be formulated. We could remain silent on whether M^\diamond is a proper or improper part of M , thus formulating a position that covers both $SOMT_2$ and $SOMT_4$:

(SOMT₅) A mental state M of a subject S is conscious iff (i) M* is a (proper) part of M, (ii) M[◇] is a (proper or improper) part of M, and (iii) M* is a representation of M[◇].

When M[◇] is a *proper* part of M, SOMT₅ accords with SOMT₄; when it is an *improper* part, SOMT₅ accords with SOMT₂. But SOMT₅ itself allows for both structures to be involved in conscious states.

Another possible view in the same ballpark is that the conscious state is a part of the awareness of it, rather than the other way round. This view may be formulated above:

(SOMT₆) A mental state M of a subject S is conscious iff S has a conscious mental state M*, such that, (i) M* is a representation of M, and (ii) M is a (proper part) of M*.

SOMT₆ appears to be defended by Kobes (1995) and Fumerton (in conversation).¹⁷

It is possible to define an umbrella view that would cover all the previous ones as specific versions. This would be done by liberally allowing both M* and M[◇] to be either a proper or an improper part of M:

(SOMT₇) A mental state M of a subject S is conscious iff (i) M* is a (proper or improper) part of M, (ii) M[◇] is a (proper or improper) part of M, and (iii) M* is a representation of M[◇].

SOMT₇ allows four different structures to qualify as conscious states: where both M^* and M^\diamond are proper parts of M (as in SOMT₄); where both are improper parts of M (as in SOMT₁); where M^* is a proper part and M^\diamond an improper part of M (as in SOMT₂); and where M^* is an improper and M^\diamond a proper part of M (as in SOMT₆).

Because SOMT₇ covers all the other versions as special cases, we may refer to it as *generic SOMT*. Its advantage is that it is less likely to turn out to be false than any other version of SOMT, since it is, in a manner, a disjunction of all these versions. Its disadvantage, however, is in its logical weakness or permissiveness: it allows many very different structures to qualify as conscious states.

3. Comparing the Plausibility of SOMT's Different Versions

In the previous section, we were concerned primarily with articulating the different versions of SOMT, not with their plausibility. In the present section, I want to make some inconclusive remarks on plausibility.

Let us distinguish between disjunctive and non-disjunctive versions of SOMT. SOMT₃, SOMT₅, and SOMT₇ are disjunctive versions, in that they consist in fact in disjunctions of the other versions. SOMT₁, SOMT₂, SOMT₄, and SOMT₆ are non-disjunctive versions.¹⁸

I have already remarked that the disjunctive versions are more plausible than the non-disjunctive ones, in that their truth requires the truth of only one of the disjuncts, but that they are correspondingly weaker, in that they claim less.

There may, however, be a significant problem with the disjunctive versions of SOMT. The term “consciousness” appears to denote a natural kind. This suggests that there is an underlying unity in the phenomenon – something

that ensures that the class of conscious states exhibits a strong homogeneity. The disjunctive versions of SOMT, in their lax construal of what makes a mental state conscious, may well miss out on this homogeneity.

As for the non-disjunctive versions, each has a stain on its plausibility. The problem facing SOMT₁ is how to account for the alleged ability of conscious states to represent themselves. Claiming that they just do is not enough. We must understand *how* this is possible. Moreover, there may be a principled problem with reconciling self-representation with naturalist accounts of mental representation.¹⁹

The main difficulty with SOMT₂ concerns the explication of the notion of logical parthood and how it might apply to *states*. Although the notion of spatial parthood, and its application to three-dimensional objects, is quite straightforward, this is not the case with logical parthood and its application to states and events.²⁰

The problem with SOMT₄ is that it appears to be only superficially, indeed *verbally*, different from HOMT. All it requires for consciousness is the compresence of a monitored state and a monitoring state. The only difference is that it calls “conscious state” not just the monitored state but the compound of both states.

SOMT₆ is implausible in a more plain and *prima facie* way. It simply appears to be unmotivated. The phenomenological facts about conscious experience do not suggest that the experience is normally part of the awareness of it, but the other way round.²¹ Moreover, if it were correct, our whole conscious life would be conducted in the second-order, as it were, since the overall conscious state would be a second-order state.

It is perhaps worth noting that one way SOMT₄ (and hence SOMT₅) could play out is as follows: the subject is (indirectly) aware of her whole conscious state *by* (or *in virtue of*) being aware (directly) of a part of it. Just as a perception

(or for that matter a painting) can represent a cabinet *by* (or *in virtue of*) representing the cabinet's front door, so a higher-order representation can represent a mental state *by* representing a part of it. In this way, M^* may represent the whole of M by representing the "other" part of M . This may be formulated as a specific version of $SOMT_4$.²²

($SOMT_8$) A mental state M of a subject S is conscious iff (i) M^* is a (proper) part of M , (ii) M^\diamond is a (proper) part of M , and (iii) M^* represents M by representing M^\diamond .

$SOMT_8$ may be used to relieve some of the implausibility attached to $SOMT_4$, in that it is different from $HOMT$ in a more substantial way.

This version does accrue a new set of problems, however. Firstly, the explication of the distinction between direct and indirect representation. Secondly, it is not clear in what way the notion of indirect representation is supposed to apply to states and events (as opposed, again, to three-dimensional objects).²³ Thirdly, what is the fact of the matter that distinguishes a direct representation of M^\diamond that serves as the basis for indirect representation of M from direct representation of M^\diamond that does not so serve? Fourthly, one may worry that what is indirectly represented is not strictly given in consciousness, and so the indirect content (if you will) of M^* is irrelevant to the structure of a conscious experience *as such*.²⁴

Another, perhaps better way to deal with the main problem facing $SOMT_4$ may be the following. There are two different ways M^* and M^\diamond may be conjoined to make up a single mental state, rather than two numerically distinct states. According to Gennaro's (1996 Ch. 2, 2002) "Wide Intrinsicity View," what makes them two parts of a single mental state is simply our decision to treat them

as such. There is no psychologically real relation between them that unites them into a single, cohesive mental state. By contrast, according to Van Gulick's (2001, 2004) "Higher-Order Global States" account and my "Cross-order information integration" model (see Kriegel 2002a, 2003b, 2005), what makes M^* and M^\diamond two parts of a single state is the fact that they are integrated and unified through a psychologically real cognitive process of information integration.²⁵ So a conscious state arises, on this view, when a mental event (M^\diamond) and the subject's awareness of it are integrated into a single unity through the relevant sort of cognitive process.

One way to capture the ontological difference between these two versions of SOMT₄ is through the mereological distinction between *sums* and *complexes* (Simons 1987 Ch. 9). A complex is a sum whose parts are essentially interconnected, or bound, in a certain way. The interconnection between these parts is an existence condition of a complex, but not of a sum.²⁶ Thus, a molecule is a complex of atoms rather than a sum of atoms, since for the atoms to constitute a molecule they *must* be interconnected in a certain way. So while for a sum to go out of existence, it is necessary that one of its parts go out of existence, this is not the case with a complex. A complex can go out of existence even when its parts persist, provided that the relationship or connection among them is destroyed.²⁷ More generally, suppose W is a whole comprised of components C_1, \dots, C_n ; then W is a sum iff W 's failure to persist entails a C_i 's failure to persist and is a complex iff its failure to persist does *not* entail a C_i 's failure to persist.

Gennaro's view seems to construe M as a mere *sum* of M^* and M^\diamond , whereas Van Gulick's and mine appear to construe it as a *complex* whose parts are M^* and M^\diamond .²⁸ This is because the latter view requires that there be a specific relationship between M^* and M^\diamond for them to constitute a conscious state, namely, the relation effected by their cognitive integration. M^* and M^\diamond would fail to

constitute a conscious state if this relationship failed to hold (or to be instantiated). There is no such provision in Gennaro's view: all it takes for M to exist is for M* and M[◇] to be compresent. This contrast can be captured through the following pair of theses:

(SOMT₉) A mental state M of a subject S is conscious iff (i) M* is a (proper) part of M, (ii) M[◇] is a (proper) part of M, (iii) M* is a representation of M[◇], and (iv) M is a *sum* of M* and M[◇].

(SOMT₁₀) A mental state M of a subject S is conscious iff (i) M* is a (proper) part of M, (ii) M[◇] is a (proper) part of M, (iii) M* is a representation of M[◇], and (iv) M is a *complex* of M* and M[◇].

These are two ontologically distinct versions of SOMT₄ (and hence of SOMT₅). The point I wish to press here is that SOMT₁₀ is substantially, not merely verbally, different from HOMT. This can be seen clearly by noting that if the monitored and monitoring states are unified through a *psychologically real* process, that process would presumably make a difference to the causal powers of the complex of the two – something that would not happen if the monitored and monitoring states are simply “summed up.”²⁹

I belabor this distinction because, unlike SOMT₉, SOMT₁₀ clearly presents a genuine – that is, substantive rather than verbal – alternative to HOMT, one that at the same time does not appeal to the problematic notion of self-representation. The problem with SOMT₉ is that there is a sense in which it retains the logical independence (postulated in HOMT) between the monitoring state and the monitored state, since it construes M* and M[◇] as completely independent of each other. This is the problem overcome through SOMT₁₀, since the latter posits an

essential connection between them. In a way, SOMT₁₀ goes a step beyond generic SOMT, in that it construes as constitutive not only the relation between M* and M but also between M* and M[◇].^{30, 31}

SOMT₁₀ does still appeal to the problematic notion of logical part. But although the notion is difficult to analyze, it is not so difficult to illustrate, as we have seen in the case of believing and being glad. That illustration suggests that there is a viable notion of logical parthood that does apply to mental states; it is just that explicating this notion is not easy. I suggest that we consider this a matter for further investigation, proceeding now on the assumption that the notion of logical parthood is sound.

Elsewhere, I have argued in greater detail for a view of consciousness that can be ontologically cast as a version of SOMT₁₀ (see Kriegel 2002a, 2003b, 2005). One thing that makes SOMT₁₀ preferable to SOMT₉ (beyond the fact that it is more clearly *substantially* different from HOMT), is that some possible cases of unconscious states appear to satisfy the condition laid out in SOMT₉. Consider, for instance, Siewert's (1998 Ch. 3) *spontaneous reflective blindsighter*, who can prompt herself to form judgements about what she blindsees, as well as reflective, second-order judgements about those judgements.³² Such a person may have an unconscious perceptual state accompanied by a second-order judgement about it. We can ascribe to such a person a state that is the sum of a first-order perceptual state and a second-order judgement about it, in accordance with SOMT₉, even though we cannot ascribe to her a conscious perceptual state. What we also cannot ascribe to her, however, is a *complex* made of the perceptual state and the second-order judgement.³³ The perceptual state and the second-order representation of it are not integrated through a cognitive process in such a way as to make the person's awareness of her perceptual state internal to that perceptual state.³⁴

It may be objected that SOMT₉ is not really a coherent position, since despite characterizing M as a mere sum of M* and M[◇], it does postulate an

essential relationship between them, namely, the relation of representation that M^* bears to M^\diamond . The objection is that a view such as Gennaro's in fact construes M as a complex, not a mere sum. However, the representational relation M^* bear to M^\diamond is essential to the identity of M^* : M^* would not be the state that it is if it did not represent M^\diamond . So if M^* did not bear the representational relation to M^\diamond , it would go out of existence. It would then fail to be the case that M 's two parts continue to exist but M itself ceased to exist – as is required for M to qualify as a complex and not a mere sum.³⁵

This brings into sharper focus the relationship R that has to hold among the parts of a whole in order to make the whole a complex rather than a mere sum. For R to be a complex-making relation, R must be (i) an existence (and identity) condition of the whole, but (ii) neither an existence condition nor an identity condition of any of the parts.³⁶ The relation between M^* and M^\diamond postulated in Van Gulick's account is of this sort, the one postulated in Gennaro's is not.

Before closing, let me note that $SOMT_8$ and $SOMT_{10}$ are perfectly compatible, and therefore can be conjoined to generate an even more specific version of $SOMT_4$:

($SOMT_{11}$) A mental state M of a subject S is conscious iff (i) M^* is a (proper) part of M , (ii) M^\diamond is a (proper) part of M , (iii) M^* represents M by representing M^\diamond , and (iv) M is a complex of M^* and M^\diamond .

Given the plausibility of $SOMT_{10}$, it appears that if the special problems attending $SOMT_8$ could somehow be neutralized, $SOMT_{11}$ would be a promising account of the ontology of conscious experience.

There are other versions of SOMT that I have not discussed at all and that do not fit comfortably into the framework I have presented in the last two sections (hence into any of SOMT₁ – SOMT₁₁). In particular, Carruthers (2000 Ch. 9) and Lurz (2003a, 2003b) have developed versions of SOMT that offer genuine and credible alternatives to the versions discussed above.³⁷ But I will not discuss their views here. My hope is that the above discussion is sufficient to bring out the special character of the kind of account envisioned by a Same-Order Monitoring approach to consciousness. I now turn to the task of arguing that SOMT has resources to deal with problems that are fatal, or at least critical, to the viability of the more traditional HOMT.

4. The Problem of Immediacy

In §2, I noted that a higher-order representation of M is a *necessary* condition for M's being conscious, because conscious states are states the subject is aware of, and awareness of something involves representation of it. I also noted that a higher-order representation of M is not a *sufficient* condition for M's being conscious, because some mental states the subject is aware of (and hence represents) are not conscious. This is why we must appeal to an "appropriate" higher-order representation of M.

From what has already been said, it is clear that some mental states the subject is aware of are conscious and some are unconscious. The question is what makes the difference between an awareness of M that guarantees M's being conscious and an awareness that does *not*. One intuitively plausible suggestion would be that awareness of M makes M conscious if it is *immediate* awareness, and that it fails to make M conscious if it is not immediate. Thus, if S is of reflective disposition, she may infer that she must be distressed or anxious about something, on the basis of how unfocused and unproductive she has been, or how

lightly she has been sleeping recently. But even if S really is distressed or anxious about something (e.g., a looming banquet with the in-laws), S's newfound awareness of it would not make the distress or anxiety conscious in the relevant sense. The reason is that the awareness lacks the requisite immediacy, being as it is mediated by reflection and inference.

So one requirement on an "appropriate" higher-order representation of M is that it make S not just aware of M, but aware of M with the requisite immediacy. The problem is that HOMT appears to fail this requirement (see Goldman 1993a, Natsoulas 1993b, Kobes 1995, Moran 2001).

Suppose S has a conscious perception of a tree. According to HOMT, the perception, M, is conscious because S has another mental state, M*, which is an appropriate higher-order representation of M. Now, surely M normally has a role in the causal process leading up to the formation of M*. Just as the tree normally has a central role in the causal process leading up to the perception of it, so the perception itself normally has a central role in the causal process leading up to the higher-order representation of it. This means that the formation of M* is not exactly simultaneous with the formation of M. Rather, there is some sort of (temporally extended) causal process starting with M and ending in the formation of M*.³⁸ This process *mediates*, in effect, the formation of M*. This, it might be argued, poses a problem for HOMT. For it appears to imply that S's awareness of her perception of the tree is mediated by the causal process in question, and is therefore *not* immediate.

David Rosenthal (1993a) addresses this problem. But before I examine Rosenthal's treatment, let me note his admission that the problem does not even arise for a view such as SOMT. Rosenthal writes (1993a: 157; italics mine):

One way to answer the question about immediacy is just to stipulate that one's being [aware] of a conscious mental state is *internal* to that state

itself. Immediacy is thus guaranteed. Our being [aware] of the state would be a *part or aspect of* the state itself, so nothing could mediate between the state and one's being [aware] of it.

The phrases “internal to” and “part or aspect of” can be understood along the lines of SOMT₁ and/or SOMT₂. They are certainly consistent with the generic SOMT₇. Since on all these versions of SOMT what makes S aware of M is M itself or a (logical) part of M, there is no causal process that mediates the formation of S's awareness of M: M *comes with* the awareness of it, if you will. The problem evaporates.

Thus generic SOMT₇ handles the problem in a relatively straightforward way. M[◇] would normally have a causal role in the process leading up to the formation of M*. But until M* is formed, the conscious state M does not exist yet. M comes into existence only upon the completion of the causal process leading up to the formation of M*. Once M comes into existence, it already envelopes within it M[◇] and M*; no further causal process is required. So M itself does not play a causal role in the process leading up to the formation M*. For M does not exist before M* does. Thus once S enters the conscious state M, S's awareness of M[◇] is not mediated in any way. In other words, once M comes into existence, no further process is needed that would mediate the formation of M*. The awareness constituted by M* is therefore immediate.

It appears, then, that SOMT faces no serious difficulty regarding the immediacy of our awareness of our conscious states. But Rosenthal claims that HOMT can account for this immediacy as well. According to Rosenthal, what is required for S's awareness of M to be immediate is not that the formation of M* *be* unmediated, but rather that it *seem* unmediated *to S*. Or perhaps even more minimally, the formation of M* must not seem mediated to S. As long as it does

not seem to S that the formation M^* is mediated, her awareness of M will be immediate. (Note that the way I am using the terms “immediate” and “unmediated,” the two are *not* synonymous, at least as applied to awareness. An awareness that is immediate may not be unmediated, as when an awareness is mediated by processes of which the subject is unaware, as we will presently see.)

There are two ways the formation of M^* may not seem mediated to S. One is when the formation of M^* *really is* unmediated. Another is when the formation of M^* is mediated, but the processes by which it is mediated are processes S is completely unaware of. If S is completely unaware of the processes that mediate the formation of M^* , M^* 's formation will seem unmediated to her, or at least it will not seem mediated to her. This latter way the formation of M^* may not seem mediated to S is the one appealed to by Rosenthal. Rosenthal's claim is that while it is true that the formation of M^* is mediated by a causal process starting with M and ending in the formation of M^* , the subject is completely unaware of this process, and therefore her awareness of M is immediate, in that it does not seem mediated to her.

To meet the requirement of immediacy, Rosenthal therefore claims that an “appropriate” higher-order representation must be *non-inferential*, where this means that the higher-order representation is not formed through a conscious inference. For such a *conscious* inference would be a mediating process of which the subject *would be* aware (since it is conscious).³⁹ In other words, where P is the process leading from M to the formation of M^* , M is conscious just in case P is unconscious; when P is conscious, M is unconscious.

(Note that the way Rosenthal uses the terms, inference is by definition conscious. To be sure, we could call certain unconscious cognitive processes “inferences,” and so allow for unconscious inference. But this is not how Rosenthal uses the term. He allows that there may be unconscious processes resembling inference in every other respect, but reserves the term “inference” to

those that are conscious. For the sake of clarity, I will align my usage with Rosenthal's. To refer to the unconscious cognitive processes that are otherwise just like inference, I will use the phrase "unconscious inferential processes."⁴⁰)

Rosenthal's treatment of the problem of immediacy may initially appear satisfactory, but it does not withstand scrutiny. The problem is to account for the difference between S's awareness of her conscious states, which is immediate, and S's (occasional) awareness of her unconscious states, which lacks the requisite immediacy. Rosenthal's suggestion is that the *conscious* states are those the awareness of which is formed through unconscious inferential processes, whereas the *unconscious* states are those the awareness of which is formed through conscious inferences.⁴¹ This suggestion, I will now argue, is unlikely to work.

Let us start by adopting a certain principle regarding inferential processes. The principle is that a conscious inference can only start from conscious "premises." More precisely, for any process P leading from mental state M₁ to the formation of mental state M₂, P is conscious only if M₁ is conscious. If M₁ is unconscious, then P is necessarily unconscious.⁴²

The problem with Rosenthal's suggestion is that M is *always* unconscious before the formation of M*, since M* is what bestows consciousness on M. So *every* process leading *from* M to the formation of M* would have to start from an unconscious state, and therefore itself be an unconscious process. This ensures that *every* higher-order representation formed through a process leading from its object (the first-order state) would be non-inferential and therefore would bestow consciousness on the first-order state.⁴³

My claim is not that there can be no awareness of M formed by conscious inference. There surely can. My claim is rather that there can be no awareness of M formed by conscious inference from M *before* M is already conscious. More generally, there can be no conscious states formed by conscious inference *from*

these states (before they are already conscious). A subject can certainly become aware that she harbors an unconscious anger at her mother on the strength of her therapist's testimony; in which case her awareness of her unconscious anger is consciously inferred from the evidence presented to her by the therapist. But in such a case, the subject's awareness of M is not formed by conscious inference *from* M (or *on the basis of* M). Rather, it is formed by conscious inference from (or on the basis of) the therapist's testimony. Rosenthal's account is incompatible with this, however, for the reason provided in the previous paragraph.

Rosenthal might modify his account of immediacy accordingly. Instead of claiming that the difference between S's awareness of her conscious states and her awareness of her unconscious states is that the former is formed through unconscious inferential processes whereas the latter is formed through conscious inferential processes, he might suggest that the former is formed through processes that do not emanate from the relevant conscious states whereas the latter is formed through processes that do.

This modified account is, however, extremely implausible, indeed somewhat absurd. On the suggestion under consideration, what makes S's awareness of M immediate is precisely that it is not formed responsively to M, but as an upshot of some other process. Whenever M happens to lead to an awareness of it, M is bound to remain unconscious. This appears to get things exactly backwards.

On the other hand, the proponent of HOMET cannot opt for the opposite modification, according to which the difference between S's awareness of her conscious states and her awareness of her unconscious states is that the former is formed through processes that *do* emanate from the relevant conscious states, whereas the latter is formed through processes that do *not*. For awareness of some conscious states may emanate from these states through conscious inferential

processes. Such inferential processes would be causal process of which S is aware, and would therefore *seem mediated* to S.

Finally, A proponent of HOMT could retreat to the view that immediacy is *not* what distinguishes the awareness we have of our conscious states from that we have of our unconscious states. But this, beside being quite *ad hoc* and *prima facie* implausible, would leave HOMT without an account of the difference between conscious and unconscious states of which we are aware. Furthermore, arguably the immediacy that characterizes our awareness of our conscious states is a phenomenon that calls for explanation regardless of its theoretical role within an account of consciousness.

In conclusion, HOMT faces a serious difficulty in its attempt to account for the immediacy that characterizes the awareness we have of our conscious states (and does not characterize the awareness we have of some of our unconscious states).⁴⁴ SOMT, by contrast, faces no serious difficulty from that quarter. In essence, SOMT's position is that the awareness we have of our conscious states is immediate simply because it *really is* unmediated.

5. The Problem(s) of Relationality

An important aspect of HOMT is the fact that it construes consciousness as a relational property: mental states are conscious in virtue of standing in a certain relation to other mental states. Many philosophers find this counter-intuitive. What it is like to be in a given conscious state seems to be an intrinsic property of the state. For some philosophers, this alone is a ground for rejecting HOMT (see Smith 1989, Gennaro 1996, Natsoulas 1999). In this section, I will argue that construing consciousness as relational is not only counter-intuitive, but also brings up the specter of two serious problems for HOMT. This would constitute a third advantage of SOMT, since these difficulties do not even arise for SOMT,

since the latter construes consciousness as an intrinsic property of the conscious state, as we will see toward the end of the section.

A decade or two ago, the most widely discussed problem in the philosophy of mind concerned the causal efficacy of mental content. After externalist accounts of content (which construe it as a relational property of mental states) became popular,⁴⁵ it was noted that this appears to render mental content causally inert.⁴⁶ The reasoning was this: only intrinsic properties of a mental state contribute to its fund of causal powers, because causation is a local affair; so if content is an extrinsic, relational property, it makes no contribution to the state's causal powers, and is therefore causally inert, or *epiphenomenal*.

That problem was never resolved to everyone's satisfaction. Different solutions, of different merits, have been offered, but no agreement is in sight.⁴⁷ One thing almost everybody accepted, though, was the thesis that the causal powers of a mental state reside fully in its intrinsic properties.^{48,49}

This thesis threatens to undermine HOMT, since the latter construes consciousness as relational. If consciousness were indeed a relational property, M's being conscious would fail to contribute anything to M's fund of causal powers. And this would make the property of being conscious epiphenomenal (see Dretske 1995a: 117 for an argument along these lines).

This is, by all appearances, a serious problem for HOMT. Why have philosophers failed to press this problem more consistently? My guess is that we are tempted to slide into a causal reading of HOMT, according to which M* *produces* the consciousness of M, by impressing upon it a certain *modification*. Such a reading does make sense of the causal efficacy of consciousness: after M* modifies M, this intrinsic modification alters M's causal powers. But of course, this is a *misreading* of HOMT. It is important to keep in mind that HOMT is a *conceptual*, not *causal*, thesis.⁵⁰ Its claim is *not* that the presence of an appropriate higher-order representation *yields*, or *gives rise to*, or *produces*, M's

being conscious. Rather, the claim is that the presence of an appropriate higher-order representation *constitutes* M's being conscious. It is not that by representing M, M* *modifies* M in such a way as to make M conscious. Rather, M's being conscious simply *consists in* its being represented by M*.

A person *could*, of course, propound HOMT as a causal thesis. But such a person would not take HOMT to be an account of consciousness itself; she would take it to be merely an account of the *causal origin* of consciousness. To the extent that HOMT is meant as an account of consciousness itself, then, it puts in jeopardy the causal efficacy of consciousness.

When proponents of HOMT have taken this problem into account, they have responded by downplaying the causal efficacy of consciousness.⁵¹ But if the intention is to bite the bullet, *downplaying* the causal efficacy is insufficient – what is needed is *nullifying* the efficacy.⁵² The charge at hand is not that HOMT may turn out to assign consciousness too little a fund of causal powers, but that it may deny it *any* causal powers. To bite the bullet, proponents of HOMT must embrace epiphenomenalism. Such epiphenomenalism can be rejected, however, both on commonsense grounds and on the ground that it violates what has come to be called *Alexander's dictum*: to be is to be causally effective.^{53,54} Surely HOMT would be better off if it could legitimately assign some causal powers to consciousness. But its construal of consciousness as a relational property makes it unclear how it might do so.

Another consequence of the alleged relationality of consciousness would be the following. According to HOMT, M's property of being conscious is just the property of being appropriately represented by another internal state. Some critics have charged that the property of being appropriately represented by another internal state is a property which internal states of inanimate objects can also instantiate (see again Dretske 1995a: 97).⁵⁵ If so, they argue, HOMT is committed to attributing conscious states to inanimate objects. Thus, when a

person harbors an appropriate representation of the internal physical state of a stone, the internal state of the stone is appropriately represented by another internal state, and so there would be no non-arbitrary way to deny consciousness to the stone's internal state.⁵⁶

Proponents of HOMT may respond that internal states can be conscious only when appropriately represented by a separate state *of the same organism* (or object). But this reply would not do. There are states of our skin that we have appropriate representations of, and yet these skin states are not conscious, even though they are states of the same organism that has the higher-order representations.

A more sophisticated rejoinder is that it need not be part of HOMT that *any* internal state can become conscious upon being appropriately represented by another internal state. In particular, it is often suggested that only *mental* states (perhaps only mental states of a certain *kind*) are such as to become conscious upon being suitably represented by another internal state.⁵⁷

This reply has less merit to it than may initially appear. Again, the problem is that we are tempted to read HOMT causally instead of conceptually. If M^* gave rise to consciousness by modifying M , then it would make a difference what characteristics M has (e.g., being mental). Thus, it could be claimed that only states with such characteristics can be so modified by being appropriately represented as to become conscious. But recall that according to HOMT, conscious states do not undergo any (intrinsic) change in response to the fact that they are appropriately represented. It is not so clear, then, what difference it makes whether an internal state has certain characteristics or not. To claim that only a certain kind of internal state is "the right kind" of state for becoming conscious upon being appropriately represented, even though nothing has to happen with those states when they are thus appropriately represented, is to introduce a completely artificial, *ad hoc* condition to the account.⁵⁸

In summary, the relational construal of consciousness lands its proponents in significant trouble. It appears to cast consciousness as causally inert and suggests consciousness may be a ubiquitous property of nature, including inanimate nature. No doubt the proponents of HOMET may devise ways of dealing with these problems. Those “ways of dealing with the problem” are likely, however, to complicate the theory considerably. And in any case, it is clearly preferable to avoid these problems altogether.

An account that construed consciousness as an intrinsic property of conscious states would therefore be preferable, inasmuch as it would not raise these problems in the first place.⁵⁹ SOMT seems to be such an account. At least the versions of it we have considered in §2 construe consciousness as intrinsic in the relevant sense.

According to SOMT₁, for instance, M is conscious in virtue of representing itself. This means that M need not stand in a relation to any *numerically distinct* state (or entity). There is a sense, of course, in which the property of being self-representing is relational, in that it is a matter of its bearer standing in a certain relation to itself. But in *that* sense, the property of being self-identical is also relational. This is clearly not the relevant sense of relationality: it is not the sense in which relationality may put in question the causal efficacy of a state, for instance. In the sense in which we are interested, a property is relational only when its bearer stands in some relation to *something other than itself*. Thus, a mental state’s property F is relational when the mental state in question instantiates F in virtue of standing in a certain relation to a numerically distinct state (or, more generally, to another entity). In *this* sense, the property of being self-representing, like the property of being self-identical, is a non-relational property.

In this regard, the part-whole relation is similar to the self-identity relation: it does not require that its bearer stand in any relation to a numerically

distinct entity. My body's property of having my left arm as a part is a non-relational property of my body, in that it does not require that my body stand in any relation to something other than itself (requiring instead that my body stand in relation to a part of itself). Likewise, if M^* is a part of M , then M 's property of standing in a certain relation to M^* (namely, the relation of being represented by M^*) is a non-relational property in the relevant sense. This ensures that in $SOMT_7$, the generic version of $SOMT$, consciousness is construed as a non-relational property in the relevant sense.

The fact that $SOMT$ construes consciousness as an intrinsic property of conscious states, whereas $HOMT$ construes it as a relational property of them, means that there are certain difficulties that arise for $HOMT$ but not for $SOMT$. This, too, is an advantage of $SOMT$ over $HOMT$.⁶⁰

6. Conclusion

For almost two decades now, the Higher-Order Monitoring Theory has been at the forefront of attempts to make the place of consciousness in nature intelligible. However, the theory faces a number of serious difficulties, some technical some fundamental. Moreover, many philosophers share the sentiment that it misses out on what is so special about consciousness. At the same time, it is built on the sound notion that conscious states are states we are somehow *aware* of being in. It is perhaps for this reason that, in recent work on consciousness, one detects an interesting, and significant, development of the monitoring approach to consciousness. A surprising number of accounts that are happy to construe consciousness in terms of monitoring attempt to bring closer together the monitoring state and the monitored state, in such a way that the two are not "independent existences," but are somehow constitutively, or "internally," or otherwise non-contingently, connected to each other.

My goal in this paper has been twofold: first, to identify this trend and lay out its conceptual foundations; and second, to suggest that the trend is indeed a positive development, in that the emerging account of consciousness can overcome a number of fundamental difficulties that have seemed to bedevil the project of the more traditional Higher-Order Monitoring Theory.⁶¹

References

- Aquila, R. 1990. "Consciousness and Higher-Order Thoughts: Two Objections." *American Philosophical Quarterly* 27 (1990): 81-87.
- Armstrong, D. M. 1968. *A Materialist Theory of the Mind*. New York: Humanities Press.
- Armstrong, D. M. 1978. *Universals and Scientific Realism*, Vol. 2. Cambridge: Cambridge UP.
- Armstrong, D. M. 1981. "What Is Consciousness?" In his *The Nature of Mind* (pp. 55-67). Reprinted in N. J. Block, O. Flanagan, and G. Guzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press/Bradford Books, 1997.
- Bell, J. L. 2004. "Whole and Part in Mathematics." *Axiomathes* 14: 284-295.
- Block, N. J., Flanagan, O., and Guzeldere, G., eds. 1997. *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press/Bradford Books.
- Brentano, F. 1874. *Psychology from Empirical Standpoint*. Ed. O. Kraus. Ed. of English edition L. L. McAlister, 1973. Translation A. C. Rancurello, D. B. Terrell, and L. L. McAlister. London: Routledge and Kegan Paul.
- Brook, A. Ms. "Consciousness and the Varieties of Externalism." Unpublished Manuscript.
- Brook, A. and P. Raymont 2005. *A Unified Theory of Consciousness*. Forthcoming.
- Burge, T. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy* 4: 73-122.
- Burge, T. 1989. "Individuation and Causation in Psychology." *Pacific Philosophical Quarterly* 70: 303-322.

- Byrne, A. 1997. "Some Like It HOT: Consciousness and Higher-Order Thoughts." *Philosophical Studies* 86: 103-129.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge UP.
- Caston, V. 2002. "Aristotle on Consciousness." *Mind* 111: 751-815.
- Chalmers, D. J. 1996. *The Conscious Mind*. Oxford and New York: Oxford UP.
- Dretske, F. I. 1993. "Conscious Experience." *Mind* 102: 263-283.
- Dretske, F. I. 1995. *Naturalizing the Mind*. Cambridge MA: MIT Press/Bradford Books.
- Fodor, J. A. 1987. *Psychosemantics*. Cambridge MA: MIT Press/Bradford Books.
- Gennaro, R. J. 1996. *Consciousness and Self-Consciousness*. Philadelphia and Amsterdam: John Benjamin Publishing Co.
- Gennaro, R. J. 2002. "Jean-Paul Sartre and the *HOT* Theory of Consciousness." *Canadian Journal of Philosophy* 32: 293-330.
- Gennaro, R. J. 2004. "Higher-Order Thoughts, Animal Consciousness, and Misrepresentation." In R. Gennaro (ed.), *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamin Publishing Co.
- Goldman, A. 1993. "Consciousness, Folk Psychology, and Cognitive Science." *Consciousness and Cognition* 2 (1993): 364-383.
- Guzeldere, G. 1995. "Is Consciousness the Perception of What Passes in One's Own Mind?" In *Conscious Experience*, ed. T. Metzinger. Padborn: Schoeningh-Verlag.
- Hossack, K. 2002. "Self-Knowledge and Consciousness." *Proceedings of the Aristotelian Society* 102 (2002): 163-181.
- Hossack, K. 2003. "Consciousness in Act and Action." *Phenomenology and the Cognitive Sciences* 2: 187-203.
- Kim, J. 1982. "Psychophysical Supervenience." *Philosophical Studies* 41: 51-70. Reprinted in his *Supervenience and Mind*, Oxford: Oxford UP.
- Kim, J. 1998. *Mind in a Physical World*. Cambridge MA: MIT Press/Bradford Books.
- Kobes, B. W. 1995. "Telic Higher-Order Thoughts and Moore's Paradox." *Philosophical Perspectives* 9: 291-312.
- Kriegel, U. 2002. "Consciousness, Permanent Self-awareness, and Higher-Order Monitoring." *Dialogue* 41 (2002): 517-540.
- Kriegel U. 2003a. "Consciousness, Higher-Order Content, and the Individuation of Vehicles." *Synthese* 134 (2003): 477-504.

- Kriegel, U. 2003b. "Consciousness as Intransitive Self-Consciousness: Two Views and an Argument." *Canadian Journal of Philosophy* 33 (2003): 103-132.
- Kriegel, U. 2004a. "Naturalizing Subjective Character." Forthcoming in *Philosophy and Phenomenological Research*.
- Kriegel, U. 2004b. "Consciousness and Self-Consciousness." *The Monist* 87 (2004): 185-209.
- Kriegel, U. 2005. "The Case Against Higher-Order Representation Theories of Consciousness." Forthcoming.
- Lehrer, K. 1996a. "Skepticism, Lucid Content, and the Metamental Loop." In A. Clark, J. Ezquerro, and J. M. Larrazabal (eds.), *Philosophy and Cognitive Science*. Dordrecht: Kluwer.
- Lehrer, K. 1996b. "Consciousness." In A. Schramm (ed.), *Philosophie in Osterreich*. Vienna: Verlag Holder-Pichler-Tempsky.
- Lehrer, K. 1997. *Self-Trust: A Study of Reason, Knowledge, and Autonomy*. Oxford: Oxford UP.
- Lehrer, K. 2002. "Self-Presentation, Representation, and the Self." *Philosophy and Phenomenological Research* 64:412-430.
- Lehrer, K. 2004. "Representation in Painting and in Consciousness." *Philosophical Studies* 117:1-14.
- Levey, S. 1997. "Coincidence and Principles of Composition." *Analysis* 57: 1-10.
- Lewis, D. C. 1991. *Parts of Classes*. Oxford: Blackwell.
- Lurz, R. 2003a. "Neither HOT nor COLD: An Alternative Account of Consciousness." *Psyche* 9 (2003).
- Lurz, R. 2003b. "Advancing the Debate between HOT and FO Accounts of Consciousness." *Journal of Philosophical Research* 28: 23-43.
- Lurz, R. 2004. "Either HOR or FOR: A False Dichotomy." In R. Gennaro (ed.), *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamin Publishing Co.
- Lycan, W. G. 1990. "Consciousness as Internal Monitoring." *Philosophical Perspectives* 9: 1-14. Reprinted in N. J. Block, O. Flanagan, and G. Guzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press/Bradford Books, 1997.

- Lycan, W. G. 1996. *Consciousness and Experience*. Cambridge, MA: MIT Press/Bradford Books.
- Lycan, W. G. 2001. "A Simple Argument for a Higher-Order Representation Theory of Consciousness." *Analysis* 61 (2001): 3-4.
- Mellor, D. H. 1978. "Conscious Belief." *Proceedings of the Aristotelian Society* 78: 87-101.
- Moran, R. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, NJ: Princeton UP.
- Mulligan, K. and B. Smith 1985. "Franz Brentano on the Ontology of Mind." *Philosophy and Phenomenological Research* 45: 627-644.
- Natsoulas, T. 1993. "What Is Wrong with Appendage Theory of Consciousness." *Philosophical Psychology* 6: 137-154.
- Natsoulas, T. 1996a. "The Case for Intrinsic Theory: I. An Introduction." *Journal of Mind and Behavior* 17 (1996): 267-286.
- Natsoulas, T. 1996b. "The Case for Intrinsic Theory: II. An Examination of a Conception of Consciousness₄ as Intrinsic, Necessary, and Concomitant." *Journal of Mind and Behavior* 17 (1996): 369-390.
- Natsoulas, T. 1999. "The Case for Intrinsic Theory: IV. An Argument from How Conscious₄ Mental-Occurrence Instances Seem." *Journal of Mind and Behavior* 20 (1999): 257-276.
- Neander, K. 1998. "The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness." *Philosophical Perspectives* 12 (1998): 411-434.
- Paul, L. A. 2002. "Logical Parts." *Noûs* 36: 578-596.
- Raymont, P. Ms. "From HOTs to Self-Representing States." Unpublished Manuscript.
- Rey, G. 1988. "A Question about Consciousness." In *Perspectives on Mind*, ed. H. Otto and J. Tueidio. Norwell: Kluwer Academic Publishers.
- Rosenthal, D. M. 1986. "Two Concepts of Consciousness." *Philosophical Studies* 94 (1986): 329-359.
- Rosenthal, D. M. 1990. "A Theory of Consciousness." ZiF Technical Report 40, Bielfield, Germany. Reprinted in N. J. Block, O. Flanagan, and G. Guzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press/Bradford Books, 1997.

- Rosenthal, D. M. 1993. "Higher-Order Thoughts and the Appendage Theory of Consciousness." *Philosophical Psychology* 6: 155-166.
- Rosenthal, D. M. 2002. "Explaining Consciousness." In D. J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford and New York: Oxford UP.
- Rosenthal, D. M. 2004. *Consciousness and Mind*. Forthcoming with Oxford UP.
- Sanford, D. 1984. "Armstrong's Theory of Perception." In R. J. Bogdan (ed.), *D. M. Armstrong*. Dordrecht: Reidel.
- Seager, W. E. 1999. *Theories of Consciousness: An Introduction and Assessment*. London and New York: Routledge.
- Shoemaker, S. 1994. "Self-knowledge and 'Inner Sense.' Lecture II: The Broad Perceptual Model." *Philosophy and Phenomenological Research* 54 (1994): 271-290.
- Siewert, C. 1998. *The Significance of Consciousness*. Princeton NJ: Princeton UP.
- Simons, P. 1987. *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- Smith, B. 1994. *Austrian Philosophy: The Legacy of Franz Brentano*. Chicago and La Salle: Open Court.
- Smith, B. and K. Mulligan 1983. "A Framework for Formal Ontology." *Topoi* 3: 73-85.
- Smith, D. W. 1986. "The Structure of (Self-)Consciousness." *Topoi* 5 (1986): 149-156.
- Smith, D. W. 1989. *The Circle of Acquaintance*. Dordrecht: Kluwer Academic Publishers. In his *Mind World*. Cambridge: Cambridge UP.
- Smith, D. W. 2004. "Return to Consciousness." In his *Mind World: Essays in Phenomenology and Ontology*. Cambridge: Cambridge UP.
- Stich, S., 1979. "Autonomous Psychology and the Belief-Desire Thesis." *Monist* 61: 571-91.
- Thomasson, A. L. 2000. "After Brentano: A One-Level Theory of Consciousness." *European Journal of Philosophy* 8 (2000): 190-209.
- Van Gulick, R. 2001. "Inward and Upward – Reflection, Introspection, and Self-Awareness." *Philosophical Topics* 28 (2001): 275-305.
- Van Gulick, R. 2004. "Higher-Order Global States (HOGS): An Alternative Higher-Order Model of Consciousness." In R. Gennaro (ed.), *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamin Publishing Co.
- Velmans, M. 1992. "Is Human Information Processing Conscious?" *Behavioral and Brain Sciences* 14: 651-669.

- Williford, K. 2003. *The Structure of Self-Consciousness*. Ph.D. Dissertation, University of Iowa.

1. This label was devised, independently, by Brook (Ms), Kriegel (2002), and Lurz (2003). It no doubt characterizes some of the accounts of consciousness I have in mind better than others, but it is the best generic label I could find.

2. In the literature on consciousness one can find a great number of arguments directed against HOMET: Aquila (1990), Byrne (1997), Caston (2002), Dretske (1993, 1995a), Goldman (1993a), Guzeldere (1995), Kriegel (2003a), Levine (2001), Lurz (2003a, 2003b), Moran (2001), Natsoulas (1002, 1993b), Neander (1998), Rey (1988), and Seager (1999) develop some of them. Some of these arguments may apply with equal force to SOMT, though some of them clearly do not.

3. Both premises 1 and 3 can certainly be denied. In particular, Dretske (1993) argues that a mental state's status as conscious does not require that its subject be aware of it. I will not discuss this issue here, as it is tangential to the main concern of the paper. For a defense of the notion that the subject necessarily has an awareness of her conscious states, see Lycan 1996, Kriegel 2004.

4. The standard example of a mental state that is higher-order represented but is still non-conscious involves a person who learns of a repressed emotion or belief through therapy and comes to represent to herself that she has the repressed emotion or belief in question, without the repressed state becoming thereby conscious. So the repressed state can remain unconscious despite being (higher-order) represented. This issue will be discussed more fully in §4.

5. The way I frame the distinction between SOMT and HOMET, the "constitutive relation requirement" is not suggested to be the *only* requirement on an appropriate higher-order representation. That is, an "appropriate" higher-order representation may be required to exhibit other features, beyond the requirement of being constitutively related to the conscious state. However, most versions of SOMT would probably see this as the key requirement for an appropriate higher-order representation.

6. In the remainder of this section, I articulate several specific versions of SOMT. The main purpose is not to evaluate these versions, but to try and articulate the conceptual foundations of this still under-discussed approach to consciousness. Hopefully, this will thereby give the reader a clearer sense of the sort of account of consciousness offered by SOMT.

7. There is one sense in which, once M is a representation of itself, it is not really a *higher-order* representation, since it is a first-order state. But in another sense, it still is a higher-order representation, since what it represents is a representation. This is, I take it, but a verbal difference, with no metaphysical significance. I will continue to use the term "higher-order representation" in this sort of context, but everything I will have to say can be said without this term.

8. We must keep in mind, however, that M's representation of itself has to be appropriate in other ways as well, in case the constitutive relation requirement is not the only requirement on appropriate higher-order representations. A similar point applies to the formulation of SOMT₂ and SOMT₃ later in the text.

9. Smith was the first to argue for this sort of view in Anglo-American philosophy. He writes (1986: 149-150): “A mental state is *conscious* if and only if it includes a certain awareness of itself, that is, in having the experience, the subject is aware of having it.” (1986: 149-150) And further along: “That inner awareness, I should like to propose, lies (to begin with) in a certain *reflexive* character in experience. That character is ascribed in the following phenomenological description, for the case of seeing a frog: ... ‘In this very experience I see this frog.’ ... Thus, the reflexive structure ‘in this very experience’ qualifies the presentational structure ‘I see this frog’: indeed, the former would seem already implicit in the latter.” (1986: 150)

10. More traditionally, this view was developed and defended by Brentano (1874) and probably also by Aristotle (see Caston 2002).

11. It is clear that the part-whole relation between M and M* would not be (or at least not primarily) a spatial or temporal part-whole relation. Moreover, it is *not* clear how such a relation would apply to states, as opposed to individual objects.

12. For discussion of the logical part-whole relation, see Lewis (1991), Paul (2002), Simons (1987), Smith and Mulligan (1983), and Mulligan and Smith (1985). A full discussion of it will take us too far afield, but it may be worthwhile to just state the logical properties of the relation of proper parthood: it is anti-reflexive (x cannot be a part of itself), anti-symmetrical (if x is a proper part of y , then y is not a proper part of x), and transitive (if x is a proper part of y and y is proper part of z , then x is a proper part of z). The relation of parthood (construed as covering improper parthood as well), by contrast, is a-reflexive, a-symmetrical, and transitive.

13. Examples of this sort are provided by Smith (1994 Ch. 3).

14. I am here working with the traditional notion of parthood, where x can be said to be a part of y even if there is no part of y that is not a part of x . In that case, x is an improper part of y , where this is more or less the same as x 's being identical with y .

15. Mereology is the theory of parts and wholes, or the part-whole relation. If the notion of a logical part-whole relation is accepted, so should the notion of logical mereology. For the legitimacy of the notion of logical mereology, see especially Paul (2002).

16. Perhaps the clearest proponent of this sort of view is Gennaro (1996, 2002). He writes (1996: 23): “We can understand consciousness as involving the property of ‘accompanied by a MET [meta-psychological thought]’ ... But we might individuate conscious states ‘widely,’ i.e., in a way that treats consciousness as an intrinsic property of those states. On this account, the MET is part of the conscious state. I will call it the ‘wide intrinsicity view,’ or WIV.”

17. Thanks to Paul Raymont for pointing out to me that Kobes’ account is really a version of SOMT₆, and to Richard Fumerton for making the case that this is a plausible view worth pausing to articulate. François Recanati (in conversation) also expressed sympathy for this sort of view.

18. To repeat, SOMT₃ is a disjunction of SOMT₁ and SOMT₂; SOMT₅ is a disjunction of SOMT₂ and SOMT₄; and SOMT₇ is a disjunction of all four non-disjunctive versions of SOMT. Note also that SOMT₃-SOMT₅ are all themselves versions of SOMT₂.

19. For development of this specific line of thought, see Levine 2001 Ch. 6, Kriegel 2003b, 2005). I will not discuss these arguments here.

20. Perhaps some help can be gotten from work on the part-whole relation in mathematics, which is clearly not a spatial relation, but rather a logical or “formal” one. For a recent treatment of the mathematical part-whole relation, see Bell 2004.

21. At least this is the case with normal conscious experiences, where the focal center of attention is on an external object, not an internal state of the subject. When one has an introspective, focal awareness of one’s internal state, the phenomenology might be captured fairly in terms of the structure suggested in SOMT₆. But this is not the case with regular, non-introspective conscious experience.

22. Note that it is also a version of SOMT₂.

23. I would like to thank Dan Zahavi for pressing me on this latter issue.

24. I would like to thank Paul Raymont for pointing this out to me.

25. Cognitive processes of integration are not unfamiliar. At the personal level, there is the conscious inference in accordance with “conjunction introduction,” as when one consciously infers that the wall is white and rectangular from one’s beliefs that the wall is white and that the wall is rectangular. At the sub-personal level, there is the widely discussed process of *binding*, as when the brain binds information from the visual cortex and from the auditory cortex to form a single, unified visuo-auditory representation of the color and sound of the same distal stimulus, say a car. On Van Gulick’s and Kriegel’s view, what makes M* and M[◇] parts of a single mental state is the fact that they are integrated into a single mental state through a cognitive process of this type. The process in question is probably different from either feature binding or conscious inference in accordance with conjunction introduction. But there is no reason to think that these are the only processes of integration employed by our cognitive system. Any process in which two separate mental states or contents are unified in such a way that they are superseded by a single mental state or content that encompasses both will qualify as a process of cognitive integration. (For a specific discussion of how such information integration may work out at the implementational level, see Kriegel 2003b.)

26. An example of a complex is the state of Hawaii (to be distinguished from the geographical location Hawaii). The state is not merely a sum of the seven islands making up Hawaii. It is also a matter of their political interconnection as answering to the same State government. If that government dissolved permanently, the state Hawaii would go out of existence, even though all its parts would persist.

27. The notion of a complex-making relation, as opposed to a mere sum, is similar to Levey’s (1997) notion of “principles of composition.” According to Levey, objects are not just sums of disparate parts, but the parts put together in accordance with a *principle of composition*.

28. At least this is how I understand Gennaro’s and Van Gulick’s views as they appear in print. It is quite possible that I am misinterpreting one or both of them. My primary interest, however, is in the views themselves, not so much in the exegesis of Gennaro and Van Gulick’s work. In particular, some passages in Gennaro’s work may suggest that he is more of a complex theorist

than a sum theorist (see especially Gennaro 1996: 29-30). More explicitly, in response to the present paper, Gennaro (this volume) argues that his view is a complex, rather than sum, one.

29. I am indebted to Paul Raymont for the crucial point concerning the difference in causal powers (or functional role).

30. The result, then, is a web of constitutive interrelations among M , M^* , and M^\diamond .

31. Moreover, $SOMT_{10}$ may help provide a fact of the matter to distinguish direct representation of a part that serves as a basis for indirect representation of the whole and one that does not. When the whole in question is a mere sum, (direct) representation of its part does not constitute (indirect) representation of it. When the whole is a complex, (direct) representation of its part does constitute (indirect) representation of it. If a cabinet could be a mere sum of its door and its frame, without the two being necessarily connected in a certain way, then representation of the door could not constitute also a representation of the whole cabinet. But since the door and the frame must be connected in a specific way in order for their whole to function in the way a cabinet does, representation of the door can double as representation of the larger unit of which the door is a part. (This may at least provide a necessary condition on doubling as indirect representation of the whole.)

32. My thanks to Terry Horgan for pointing me to this example.

33. As Siewert notes – though not in so many words – we *can* ascribe to her a complex of the first-order judgment and the second-order judgment; but we still cannot ascribe to her a complex of the first-order perceptual state and the second-order judgment.

34. Gennaro's particular version of $SOMT_9$ is a bit more complicated and compounds other implausibilities. Thus, according to Gennaro M^* is an *unconscious part* of the conscious state that M is. This is doubly implausible. First, although mental states are bearers of the property of being conscious, it is not clear in what sense state-parts can be said to be conscious or unconscious; and second, even if there was a sense in which state-parts could be said to be conscious, presumably what would make a state-part conscious is that it is part of a conscious state – so the notion of an unconscious part of a conscious state would be contradictory.

35. A similar objection may be that Gennaro's view, in order to be at all plausible, must require that M^* and M^\diamond be roughly simultaneous and occur in the same subject's head, but such relations would make his view a complex view rather than a sum view. In response, it may be claimed that temporal and spatial relations are not substantive enough to be complex-making.

36. If we take into account the point raised in the previous endnote, we must also require (iii) that R not be a merely temporal or spatial relation.

37. According to Carruthers, M^* is somehow inherent in M in virtue of the fact that it is part of M 's inferential role in S 's cognitive system that it is disposed to bring about the occurrence of an M^* -type state. This inferential role determines the content of M , therefore M^* is a determinant of M 's content. According to Lurz, M^* represents not M itself, but rather M 's content. It is the fact that M not only represents what it does, but is also accompanied by a representation of what it represents, that makes M conscious. Lurz explicitly calls his view "Same-Order Monitoring."

38. There are places where Rosenthal claims explicitly that there is normally *no* causal connection between M and M* (e.g., Rosenthal 1990: 744). These comments are sporadic and unmotivated, however. The resulting HOMT is, if anything, less plausible than it should be (see Kobes 1995).

39. The reason the subject would necessarily be aware of this process is that it is conscious, and conscious states and processes are states and processes the subject is aware of having or undergoing.

40. It is important to stress that no substantive issue is at stake here, only a verbal one. If we insist that there are unconscious inferences, Rosenthal would only need to rephrase his thesis. Instead of claiming that M* is an appropriate representation of M only if it is non-inferential, he could claim that that M* is an appropriate representation of M only if it is non-schminferential, where “schminference” is a conscious inference.

41. One might interpret the view otherwise, though. The suggestion might be thought to be that M* is not formed through any process, but rather “forms” somewhat simultaneously – or that it is formed either through an unconscious inferential process or through no process whatsoever. However, the notion of a mental states that is unformed, or forms spontaneously through no process, is not obviously intelligible.

42. It is important to distinguish here between a process being conscious and the process’ product being conscious. There are certainly inferential processes whose product is conscious even though the “premises” are not. But that is not the same as the inferential processes being themselves conscious.

43. Consider a normal case in which a higher-order representation M* is formed. Before M* is formed, M is not conscious (since it is not represented). There then takes place a process leading from M to the formation of M*, at the end of which M becomes conscious (due its representation by M*). What Rosenthal must do is distinguish between processes that would make M* an immediate awareness of M and processes that would make M* an awareness lacking the requisite immediacy. His suggestion is that the former are unconscious inferential processes, whereas the latter are processes of conscious inference. However, at the beginning of all these processes, M is supposed to be unconscious. So if we accept the principle that conscious inference can only start from conscious “premises,” the fact that at the outset of the process M is unconscious means that the process cannot possibly be a conscious inference. So in fact no awareness of M can be formed through a conscious inference from M (before M is already conscious). (It is, of course, possible to make a conscious inference from one of one’s conscious state to an awareness of that state. But the awareness formed through such inference is not the kind that initially bestows on the state its consciousness, since the state must already be conscious for the awareness of it to arise in this way.) Therefore, there is no explanatory force in the distinction between awareness of M formed by conscious inference from M and awareness of M formed by unconscious inferential processes emanating from M. It is not this distinction that marks the difference between immediate awareness of M and awareness of M that is not immediate.

44. Another, related problem with Rosenthal’s original suggestion for distinguishing immediate awareness from awareness lacking immediacy – which I did not discuss in the main text – is brought up by Kobes (1995: 293): “...suppose that, by feedback training or neurosurgery, I become [aware] of the...inference that yields the HOT [higher-order thought M*]. Then it follows

on Rosenthal's view that [M] is no longer conscious. But it is not credible that the addition of [awareness] of processes whereby the HOT is derived should cause loss of consciousness of the first-order state." That is, it is absurd to think that a mental state that is conscious would suddenly become unconscious when the subject suddenly becomes aware of the process that mediated the formation of the higher-order representation.

45. See Putnam 1975, Burge 1979.

46. Perhaps the most poignant presentation of the problem is Stich's (1979).

47. Perhaps the most common approach was to claim that even if mental content lacks causal powers, it is nonetheless *explanatorily* relevant in psychology (see Burge 1989). Another popular strategy, identified with the internalist camp, was to construct a notion of *narrow content* – that is, content which is fully determined by the intrinsic properties of the state whose content it is (see Fodor 1987) – and to claim that this narrow content is the causally efficacious content.

48. For a defense of this thesis in this context, see Kim 1982.

49. Sometimes, it has been claimed not that causal efficacy resides solely in intrinsic properties, but that it resides solely in properties that *supervene* on intrinsic properties ("locally supervenient" properties). This does not make a difference to the present argument, though. The present argument is based on the fact that HOMT construes consciousness as an extrinsic relational property. But HOMT *also* construes consciousness as not locally supervenient. Thus, according to HOMT, two mental states that are intrinsically indistinguishable can differ in consciousness: one is conscious and one is not (because one is appropriately represented and one is not). If so, the property of being conscious is not locally supervenient.

50. By calling the thesis "conceptual," I do not mean to suggest that it is merely a piece of conceptual analysis of the *concept* of consciousness. Rather, I mean to say that the thesis in question is not a thesis about what causes or brings about consciousness, but about what consciousness *is*. Also, by saying that the thesis is not causal, I don't mean to suggest that causality has no place in the HOMT view. Rather, I want to emphasize that the view is not that suitable higher-order monitoring *causes* consciousness, but rather that it *constitutes* consciousness.

51. Thus Rosenthal (2002: 416; italics mine): "It's easy to *overestimate* the degree to which a state's being conscious does actually play any [causal] role... [In fact,] whether or not a state is conscious will not affect the state's [causal] role in planning and reasoning."

52. Epiphenomenalism about Consciousness has been explicitly propounded by some (Velmans 1992, and to a significant extent, Chalmers 1996). But I take it that it is still a virtue of an account of consciousness that it does not render consciousness epiphenomenal. Epiphenomenalism is a liability, not an attractive feature.

53. Kim (1998) is responsible for reintroducing this dictum into philosophical discourse. In fact, what HOMT violates is an even weaker principle: to be is *at least* to be causally efficacious.

54. There may also be an epistemological problem involved in epiphenomenalism: if genuine knowledge requires causal interaction, as some philosophers have maintained (e.g. Goldman 1967), there can be no knowledge of epiphenomenal entities or phenomena. This would make

HOMT entail the absurdity that we cannot, in principle, have any knowledge of the existence of consciousness.

55. Dretske (1995a: 97) writes: “Some people have cancer and they are conscious of having it. Others have it, but are not conscious of having it. Are there, then, two forms of cancer: conscious and unconscious cancer?... Experiences are, in this respect, like cancers. Some of them we are conscious of having. Others we are not. But the difference is not a difference in the experience. It is a difference in the experiencer – a difference in what the person knows about the experience he or she is having.” See also Van Gulick 2001.

56. Since this reasoning applies to just about everything in nature, some have framed the problem in terms of panpsychism: HOMT appears to lead to panpsychism according to which anything in nature is capable of consciousness. Although some philosophers have flirted with panpsychism (e.g., Chalmers 1996), such panpsychism would not sit well with the reductive and demystifying ambitions of HOMT.

57. Thus Lycan (1990a: 758-9; I am quoting from the reprint in Block et al. 1997): “What is it that is so special about physical states of that certain sort, that consciousness of them makes them ‘conscious’? That they are themselves mental... It seems psychological states are called ‘conscious’ states when we are conscious of them, but nonpsychological things are not.” Lycan’s view is particularly implausible, as he seems to hold that there is nothing substantially different about mental states that makes them conscious upon being suitably represented – it is simply that we are unwilling to *call* internal states of inanimate objects conscious when they are suitably represented.

58. After all, as we can see with Lycan’s view (see the previous footnote), there is nothing theoretically (or explanatorily) relevant in the fact that these states are mental. The upshot must be that there is an arbitrary fact which makes suitably represented mental states, but not other suitably represented internal states, conscious. In Lycan’s case the arbitrary fact in question is the fact that we are willing to *call* the former, but not the latter, “conscious.” This line of rejoinder, if seriously pursued, would be at odds with the fact that conscious states most probably constitute a *natural kind*, and in any event seem to share something objective that is common and peculiar to them.

59. Or if it would, this would have to be for some other reason.

60. This is also connected to one of HOMT’s best-known difficulties, namely, the problem it faces with second-order misrepresentations of the very existence of a first-order mental state (see Byrne 1997, Neander 1998, Seager 1999 Ch. 3, Levine 2001 Ch. 4). As several authors have noted, this problem does not present itself for a view such as SOMT₁: a mental state may misrepresent its own *properties*, but it is impossible that it should misrepresent its own *existence* (see Caston 2002, Kriegel 2003a, Raymont Ms). Whether this solution extends to other versions of SOMT is something we will not have occasion to consider here.

61. For comments on earlier drafts of the paper, I would like to thank Rocco Gennaro, Terry Horgan, Joe Levine, Robert Lurz, Barry Smith, Cybele Tom, Dan Zahavi, and especially Paul Raymont. I have also benefited from presenting this paper, or parts of it, to audiences at Cambridge University, Dartmouth College, Institut Jean Nicod, North Carolina State University, and University at Buffalo. I would like to thank the audiences there as well.