

# Higher-Order Thoughts and the Unity of Consciousness

Lowell Friesen  
Booth University College

Rosenthal is perhaps best known for his higher-order thought theory of consciousness, but he has also expanded his theory to account for the unity of consciousness. His account posits two distinguishable mental mechanisms. I argue that, although both mechanisms may serve to unify consciousness in certain ways or to some degree, they are not sufficient to account for all of the different ways in which consciousness is unified. Thus, Rosenthal's account fails as a general account of conscious unity.

Most of the time our attentions are directed outwardly at the world around us and not on our inner, experiential lives. Our conscious experiences are ever present and, for many, a mundane feature of existence. But conscious experience is a rich and multi-faceted phenomenon. For the curious and reflective, it is a seemingly limitless object of study. In recent decades, we have been presented with numerous examples of subjects whose experiential lives appear to be markedly different than our own. When we contemplate what it might be like to be one of these subjects, it becomes apparent that our own experience is unified in ways that we were previously unaware of or deemed entirely unremarkable. And once we suspect that certain features of our experience may not be ubiquitous, we look for an explanation. What is responsible for the unity of consciousness?

A fair bit of work has been done cataloguing the different kinds of conscious unity, but work on attempting to uncover the mental mechanisms responsible for them is still in the early stages. Rosenthal is most known for his higher-order thought (HOT) theory of consciousness, but he has also extended his theory to account for conscious unity. This essay is a critical examination of Rosenthal's account. I will argue that it fails as a general account of conscious unity. The discussion will focus on Rosenthal's account, but it will have broader implications. One need not be a higher-order theorist to be tempted to use elements of Rosenthal's account to explain conscious unity. Many of my criticisms will apply whether those elements appear within the context of a higher-order theory of consciousness or not.

The essay will begin with a brief discussion of some different kinds of conscious unity and a presentation Rosenthal's account. The critical portion of the essay will begin

with a discussion of some general, but important, questions Rosenthal fails to address. It will then turn to an examination of the two mental mechanisms Rosenthal invokes to explain conscious unity and it will argue that there are some kinds of conscious unity for which neither mechanism is able to account. The essay will conclude with a brief discussion about how Rosenthal's account of conscious unity presupposes the same kind of representational mismatches that some have argued pose a serious problem for higher-order theories of consciousness.

## The Unity of Consciousness and its Many Forms

Reflection on the nature and structure of conscious experience reveals that experience is unified in a number of different ways. Some forms of conscious unity apply to simultaneous (synchronic) conscious states. Others apply to states that are conscious at different times (diachronic). Perhaps the most fundamental kind of synchronic unity is *phenomenal unity*. A number of experiential states are phenomenally unified when they are experienced together as part of a single overarching or global experience at a time Tye (2003). My current experience, for example, is a complex one. I have a visual experience of my computer monitor. I hear music playing in the background. I can feel the chair pressing into my back and the keyboard beneath my fingers. I can taste the remnants of my last sip of coffee. Each of these experiences are independent of each other in the sense that it would be possible to have any one of them without the others. To use Nagel's (1974) well-known terminology, there is "something it is like" to see my computer monitor, and there is "something it is like" to hear music playing. But there is also "something it is like" to see my computer monitor while hearing music. The latter experience encompasses the former. When a number of experiences are encompassed by a larger experience in this way, they have a conjoint phenomenology. Theorists have used a variety of terminology to refer to this relation. Parfit (1984), Lockwood (1989), and Dainton (2000), for instance, say that such experiences

---

Please cite the published version of this essay: *The Journal of Mind and Behavior*, Autumn 2014, Volume 35, Number 4, Pages 201–224.

Journal website: <http://umaine.edu/jmb/>

are “co-conscious.” Bayne and Chalmers (2003) and Bayne (2008, 2010) say that such experiences are “subsumed” by the experience of which they are both a part.

Another kind of synchronic unity is *spatial unity*. Under normal circumstances, and with correctly functioning sensory systems, the objects we perceive via our senses are given to us as spatially related to each other. They appear as located within a three-dimensional phenomenal space. Kant famously argued that this kind of spatial integration is a necessary feature of experience. Dainton (2000, ch. 3), drawing on some scenarios described by Dennett (1978), has challenged this claim. Suppose a team of scientists gain complete control over my sensory inputs. Suppose they give me a visual and auditory experience of a beautiful mountain vista accompanied by birds singing in some nearby trees. And suppose they give me a tactile and somato-sensory experience of swimming under water. What would the nature of my experience be if I were then given the sensation of scraping my shin against a sharp underwater rock? It is conceivable that the rock and the mountain would not be perceived as inhabiting a single three-dimensional spatial field and, hence, as not being spatially related to each other. A less fanciful example involves a common experience among the tall buildings in large urban centres. I often find that, upon first hearing the siren of an emergency vehicle, I cannot tell where it is coming from. Eventually its location becomes clear—or at least it becomes clear where the sound *seems* to be coming from—but there is a period of time where its location relative to other objects in my sensory field is decidedly unknown.

A third kind of unity, *object-unity*, involves the way we perceive objects (Tye, 2003; Bayne & Chalmers, 2003). As I look at the coffee mug sitting on my desk, I perceive redness and cylindricality. What is noteworthy about my experience is that it is of a single object in which cylindricality and redness both inhere. I do not, for instance, experience redness as belonging to one object and cylindricality as belonging to another, even though it is possible to have an experience in which this is the case—imagine gazing at a red cube beside a blue cylinder. What is remarkable about this kind of experience is that the different visual properties of objects are detected by different parts of our visual system and in different regions of the brain before coming together in a single visual experience of a complex object. Attempts to explain this phenomenon have come to be known as solutions to the binding problem (Roskies, 1999; Treisman, 1999). The literature on this subject contains many examples of binding failures, many of which involve “illusory conjunctions” of the colors and shapes of the objects in a stimulus. These illusory conjunctions have been elicited in normal subjects (Treisman & Schmidt, 1982; Treisman, 1998), but there are also documented cases involving subjects suffering from certain kinds of brain damage who experience these illusory conjunctions in a more persistent way (Coslett & Lie, 2008; Treisman,

1998; Ward, Danziger, Owen, & Rafal, 2002). The binding problem, as it is commonly known, concerns the visual system and visual perception. As such it concerns a kind of *intra-modal object-unity*. But our experiences of objects often draw from a number of different sensory modalities. If I hold my hand near the mug, I feel its warmth. If I pick it up, I sense its weight. I perceive not only cylindricality and redness as inhering in it, but also warmth and heaviness. Thus, object-unity can also be *inter-modal*.<sup>1</sup>

Phenomenal unity and spatial unity are paradigm examples of synchronic conscious unity. A paradigmatic kind of diachronic unity is *stream-unity*. My current global experience is part of a temporally extended episode of consciousness that began when I awoke this morning. The same is true for each of the momentary conscious states I enjoyed along the way. Together, they form a stream of consciousness. Whenever a pair of momentary global experiences are part of the same temporally extended episode of consciousness in this way, they are stream-unified.

A slightly different kind of diachronic unity involves our ability to ascribe our own mental states to ourselves. For instance, I can think of the experience I have as I reach for my coffee mug as my own. I can do the same for experiences I had the past. Any pair of experiences that I self-ascribe in this way are *subject-unified*. There is a vast literature, both in philosophy and psychology, about the ability to self-ascribe mental states and the relation of this ability to self-knowledge and personal identity. Inevitably, these discussions wade into metaphysical and epistemological questions about the nature of the “self” and of what a subject of experience is. There are nearly as many opinions about these questions as there are theorists writing about them. I do not want to take up these disputes here except to say that the concept of a subject of experience that I will be working with in the following pages is relatively minimalist in nature. Any being that has conscious mental states counts as a subject of experience. And whether or not that individual has the capacity to self-ascribe mental states will depend on having certain concepts and cognitive capacities. An individual may have the ability to self-ascribe a mental state even if it lacks, in colloquial terms, a “personal identity” or the ability to tell any sort of social and personal narrative about herself. There are additional kinds of conscious unity, but this list will provide an adequate place to start from.

<sup>1</sup>My object-unified experience with respect to my coffee mug is a kind of synchronic unity. But it seems as though object-unity can also be a diachronic unity, depending on the nature of the perceived object. A melody, for instance, is a temporally extended perceptual object; in order for a number of notes to be heard as a melody they must be heard in succession. If it is coherent to think of a melody as a perceptual object, then our experience of musical melodies involves a kind of diachronic object-unity.

### Rosenthal's Account of Conscious Unity

Rosenthal's (2003, 2005c) account of conscious unity is an extension of his higher-order theory consciousness (1986, 1997). Higher-order theories of consciousness make conscious experience a matter of representation; a mental state,  $m_1$ , becomes a conscious mental state when its subject is in a mental state,  $m_2$ , that represents  $m_1$  in the appropriate way. Some, like Carruthers (1996, 2000) and Rosenthal, maintain that the representing (higher-order) state,  $m_2$ , must be thought-like in nature. Others, like Armstrong (1968, 1981) and Lycan (1987, 1996), maintain that  $m_2$  must be perception-like in nature. Most higher-order theories maintain that the (lower-order) represented state,  $m_1$ , and  $m_2$  must be numerically distinct. Kriegel (2009), however, has defended a one-state higher-order theory according to which a mental state becomes conscious when it represents *itself* in the appropriate way (i.e.,  $m_1$  and  $m_2$  are the same state). Another kind of higher-order theory, defended by Brown (2014), eschews the relational aspect of traditional higher-order theories and posits that conscious experience is just the product of being in a mental state that has higher-order content. Rosenthal's hot theory, as a general theory of conscious experience, will be considered in the final section. Here my focus will be on the elements of his theory that are important for his account of conscious unity.

At any given time during our wakeful hours we are in numerous conscious states. Every conscious state is accompanied by a hot that targets it. The atomistic nature of this theory, says Rosenthal, "may seem to prevent it from being able to explain our sense of the unity of consciousness. If each conscious state owes its consciousness to a distinct hot, how could we come to have a sense of such unity?" (2005c, p. 340). How is it that "all of our conscious states seem to belong to a single, unifying self?" (p. 340).

Rosenthal answers this question by maintaining that hots "operate on many of our mental states not singly, but in large bunches" (2000, p. 226). They target and represent many different mental states all at once. Consider the well-known cocktail party effect. Cocktail parties are often noisy, with numerous conversations going on all at once. In spite of this, the mention of one's name in a conversation across the room is often enough to draw one's attention to it. If one's name were mentioned in a different conversation, it would have the same attention-grabbing effect. But we do not hear the many conversations as articulated conversations. Rather, we hear them as a background din. What this suggests, according to Rosenthal, is that "one's hots group many auditory sensations together, making them conscious only as an unarticulated bunch" (2000, p. 227). A single hot can make a number of mental states conscious all at once, and conscious lower-order states become unified when they are all represented together by a single hot.

Co-representation at the higher-order level, however, can-

not account for conscious unity by itself.

Wholesale operation of hots...doubtless helps to induce some conscious sense of unity among our mental states. But that will only go so far. Since no single hot covers all our conscious states, the basic problem remains. How can we explain a sense of unity that encompasses states made conscious by distinct hots? (2005c, p. 342)

Co-representation at the higher-order level cannot account for the sense of conscious unity by itself, because if no single hot represents all of one's conscious lower-order states at once, there will always be at least one pair of lower-order states that are not subsumed by the same hot.

In response to this problem, Rosenthal presents another feature of hots that serves to unify experience. A hot is a thought that ascribes a mental state to an individual. It is a thought with a content of the form "I am in such and such a mental state." It ascribes a mental state to an individual by referring to the subject of that mental state via the reflexive indexical "I." This gives the owner of the hot the sense that she is directly aware of the subject of the hot's target state. Because hots refer to the subjects of their target states indexically, their contents are "bare" with respect to the way in which they characterize those subjects. Hots do not refer to those subjects via descriptions. This means that hots do not differentiate the bearers of their target states in any way. Nothing about the mental "I" by itself gives any information about its referent. So nothing in the content of a hot indicates, by itself, that the bearer of its target state is the same as or different than the bearer of another hot's target state. This, according to Rosenthal, facilitates a certain kind of subjective impression in us. "[O]ur seeming to be aware in a direct and unmediated way of the self each hot refers to tilts things towards apparent unity. Since we seem to be directly aware of the self in each case, it seems subjectively as though there is a single self to which all of one's hots refer, a single bearer for all our conscious states" (2005c, p. 344).

Thus, Rosenthal offers a kind of two-pronged account of conscious unity: conscious unity is due in part to the fact that hots often co-represent multiple lower-order states at once and in part to the fact that they seem to ascribe their target states to a common subject. In what follows, I will refer to these two aspects of Rosenthal's account as the *co-representation mechanism* and the *common-ascription mechanism* respectively.

### Refining Rosenthal's Account

Rosenthal's account of conscious unity gives rise to several immediate questions. The first has to do with the fact that there are many different kinds of conscious unity but only two mental mechanisms to do the unifying work. How

are we supposed to understand the explanatory domains of these two mechanisms? The second has to do with the co-representation mechanism. What, precisely, is the content of a *NOT* when it collectively represents a number of lower-order states? The third has to do with the common-ascription mechanism. Can direct reference by itself serve as a unifying mechanism? None of these questions raises insurmountable problems for Rosenthal's account, but they reveal important issues that he does not address directly.

### *Unities and Mechanisms*

Rosenthal introduces a few kinds of conscious unity, but when the discussion turns to offering an account of conscious unity the plurality of conscious unities disappears from the picture. One is left with the sense that the explanandum is a single homogeneous property and that it is conscious unity *per se* that is being explained rather than one or other specific kind of unity. One is also left with the sense that the explanans, i.e., the co-representation mechanism and the common-ascription mechanism, work in a cooperative fashion to produce conscious unity. These ways of conceiving the explanandum and the explanans are both problematic.

*Conscious unity as a homogeneous mental property.* Consider, first, the explanandum. The different forms of conscious unity vary greatly in nature and character. Perhaps the most significant structural difference between them is that some of them are synchronic and others diachronic. Whichever mental mechanisms produce the several kinds of synchronic unity, their effect on a subject's overall experience at a time is profound. For instance, if a pair of mental states are phenomenally unified, they are part of the same global experience. If not, then not. Nothing, it would seem, would have a bigger impact on what a subject's global experience is like at a time than which mental states are part of that experience. The phenomenal impact of object-unity is almost as dramatic. I am currently experiencing my coffee mug as cylindrical in shape, red in colour, and warm to the touch. If my sensory systems were not functioning properly and I did not experience redness as belonging to the object that is cylindrical and currently warm to touch, my present experience would be very different. The mechanisms responsible for the different diachronic unities, on the other hand, seem to have a comparatively small impact on the phenomenology of one's experience at a time. For instance, the first visual experiences I had this morning are stream-unified with the visual experiences I have now. And even though these visual experiences are stream-unified, the fact that they are stream-unified has little or no bearing on the overall character of my present experience. I could have had very different visual experiences when I woke and have had qualitatively identical visual experiences now. Or consider subject-unity. If I ascribe a (past or present) mental state to myself, I introduce a phenomenological component to my present experience—

having a conscious thought is phenomenologically different than having the same thought unconsciously—but having a conscious thought does not dramatically effect the overall character of my experience. Synchronic and diachronic unities vary greatly with respect to their impact on the phenomenal character of present experience.

These differences between the kinds of synchronic and diachronic unity are important for two reasons. First, they reinforce the point that conscious unity is not a homogeneous mental property. Conceiving of the explanandum of a theory of conscious unity in this way is misguided. Second, and more important, these differences make it *prima facie* doubtful that only two mental mechanisms are responsible for them all. As the above discussion has made clear, some kinds of conscious unity have a much greater effect on the overall phenomenal character of one's experience at a time than others do. This already suggests different mechanisms. What also suggests differences in mechanism is that these different kinds of unity are conceptually independent of each other and, in some cases, have been empirically demonstrated to be independent of each other. For instance, one kind of unity (phenomenal unity) has to do with how all of the phenomenal properties one experiences at a given time are experienced together in a single encompassing experience. Another kind of conscious unity (object-unity) has to do with how a variety of different phenomenal properties get bound together into a single percept of a complex object. As was pointed out above when the concept of object-unity was introduced, research on the binding problem has produced numerous documented failures of object-unity. But these failures of object-unity seem to occur without producing simultaneous failures of phenomenal unity. Likewise, failures of object-unity can occur without failures with respect to subject-unity. Binding failures, for instance, can occur without those experiencing the failures also attributing their mental states to different subjects. Of course, merely pointing to the many ways that different kinds of conscious unity seem to be able to come apart does not demonstrate that it is *impossible* for one or two mental mechanisms to be responsible for all of them, but it does introduce doubt about the adequacy of any account like Rosenthal's that seems to posit a one-many ratio of mechanisms to unities.

Rosenthal has at least one theoretical resource at his disposal that may help to blunt this worry. The co-representation mechanism contributes to conscious unity by representing multiple lower-order states in a single higher-order representation. Rosenthal is largely silent about the details of this kind of collective higher-order representation, but there is nothing in principle that should prevent him from exploiting collective representation in different ways. He could maintain, for instance, that it is not only the fact that multiple lower-order states can be co-represented by the same *NOT* that contributes to conscious unity but also the *way* in which

they are co-represented. Later I will argue that mere co-representation is not enough to account for conscious unity, but that the higher-order representations must also represent certain relations between lower-order states and their contents. Nothing prevents Rosenthal from accepting this claim. He could maintain that conscious states are unified in one way in virtue of a HOT representing them as related in a certain way, while conscious states are unified in another way by virtue of a HOT representing them as related in a different way. This would also be consistent with the way in which Rosenthal conceives of the function of HOTS in other writings. HOTS, he says, often “play a partially interpretive role” (2005b, p. 211). Dental patients, for instance, can experience the vibrations of the drill together with their own fear as pain. If they are told that the relevant nerves have been anaesthetized, this can often change their experience.

It is doubtful that experiencing vibration and fear as pain is just a matter of a HOT representing certain *relations* between the lower-order sensory states. But what this shows is that Rosenthal grants HOTS a significant degree of representational license, and this may provide Rosenthal the theoretical space within which to develop an adequate response to the above worry. This may be a way for Rosenthal to address the apparent disparity between the number of different kinds of conscious unity and the number of mental mechanisms available to account for them, but for the moment I will set it aside and consider another worry about the way in which the mental mechanisms seem to be presented in his account.

*Cooperating mental mechanisms.* What is clear from Rosenthal’s account is that the co-representation mechanism and the common-ascription mechanism each have a role to play in producing conscious unity. What is less clear, however, is exactly how the two mechanisms work together or play off of each other to perform this role. What are their respective explanatory domains? Are their explanatory domains completely distinct so that each mental mechanism has its own species of conscious unity to account for? Or do their explanatory domains overlap partially? That is, are there some kinds of conscious unity for which both mechanisms are together responsible and others for which one or other mechanism independently accounts for? Or, finally, do their explanatory domains overlap completely? Although Rosenthal does not address these questions explicitly, it is possible to read him as endorsing the view that the two mental mechanisms have largely, if not completely, overlapping explanatory domains. There are a number of problems with this view.

First, just as the different species of conscious unity vary greatly in nature and character, so do the mental mechanisms of co-representation and common-ascription. Though both crucially involve the representational contents of higher-order states, they are associated with different aspects of those higher-order contents. In the latter, unity is achieved

as a result of the way in which higher-order states refer to the individuals to which their target lower-order states are ascribed. In the former, unity is achieved as a result of the fact that higher-order states represent the contents of their target states. Referring to an individual is a very different kind of thing than characterizing the contents of mental states. There is no antecedent reason to believe that these very different aspects of HOTS and their contents would be associated with the same or similar unifying functions. Thus, there is reason to be suspicious of an account according to which the different mental mechanisms involved in unifying consciousness would each be partially responsible for each of the many different species of unity.

Second, if the explanatory domains of these two mechanisms overlap in this way, additional questions arise about how this is to be understood. Select any arbitrary kind of conscious unity. Are both mechanisms *necessary* for it to obtain? If one of the mental mechanisms were absent, would the conscious states fail to be unified in the relevant way? Or conversely, is the relevant kind of conscious unity causally overdetermined by the two mental mechanisms such that each mental mechanism is *sufficient* to produce the relevant kind of conscious unity on its own independently of the other, entailing a kind of causal redundancy?

Both of these scenarios are problematic. Suppose both mechanisms are necessary to produce a given kind of unity. We can then ask what it would be like for a subject in a situation in which only one of the mental mechanisms was operational? Would the subject’s experience be merely partially unified in the relevant way or would it fail to be unified at all? If the latter, we are left to wonder what the causal role of the remaining mental mechanism is. If the former, can we make sense of a scenario in which a given sort of conscious unity only partially obtains? I will argue that we cannot. For many of the different kinds of conscious unity we have been presented with, the notion that conscious states can be partially unified along any of these different dimensions seems incoherent.

Take phenomenal unity as an example. Is there any sense that can be made of situation in which a number of conscious states are partially phenomenally unified? This question has been debated extensively in the literature on split-brain patients. Split-brain patients are individuals who have had their corpus callosum surgically severed. The experimental data suggest that the two hemispheres of the upper brain in these patients do not receive the same sets of sensory information. For instance, if an object is shown to the left hemisphere, the patient is able to identify it verbally—language is a left hemisphere function—but not manually with the left hand—motor control of the left hand is a right hemisphere function. The opposite happens if the object is shown to the right hemisphere. This has led to speculation about what the experiential lives of these patients are like. Dainton (2000) and Bayne

(2008, 2010) argue that they have a single experiential field or stream of consciousness. Others, like Sperry (1968, 1984), Puccetti (1981), Marks (1981), Tye (2003), Koch (2004), and Schechter (2012) express varying degrees of sympathy for the view that they have two (non-overlapping) experiential fields. On this view, there may be a large degree of qualitative similarity between the two fields at any given time, but none of the conscious states in one experiential field or sphere is token-identical to any conscious state in the other. Thus, neither of these views interprets the overall phenomenal field of a split-brain patient as being partially unified.

To my knowledge, Lockwood (1989, 1994) is the only theorist to defend the view that split-brain patients have two partially overlapping fields of experience. This amounts to the claim that it is possible for a pair of conscious states,  $c_1$  and  $c_2$ , to be parts of single global experience and for the pair,  $c_2$  and  $c_3$ , to be parts of a single global experience without, at the same time, the pair,  $c_1$  and  $c_3$ , being parts of the *same* global experience. The main criticism of Lockwood's view has been that it seems to be incompatible with any plausible way of understanding what it means for a pair of conscious states to be part of the same global experience or for them to have a conjoint phenomenology (Bayne, 2008, 2010; Dainton, 2000). The relation of having a conjoint phenomenology is a phenomenal relation. The only access we have to it is through our own experience, and it is extremely difficult, if not impossible, to conceive of a scenario in which the relation is instantiated in the way Lockwood suggests. Although Lockwood has gone to some length to defend the partial-unity interpretation of the split-brain data, he has himself expressed doubts about its coherence (1994, p. 95). If we cannot make sense of a situation in which conscious states are partially phenomenally unified, then we have reason to be skeptical of any account that suggests phenomenal unity is produced via a pair of mental mechanisms such that one of the mechanisms would produce a partially phenomenally unified experience in the absence of the second mechanism.

Dividing the unifying labour between a pair of mental mechanisms in this way is problematic with respect to phenomenal unity. What about some other kinds of unity? Subject-unity, for instance, also appears to be an all-or-nothing phenomenon. What could it mean for a pair of mental states to be partially subject-unified? It is certainly possible for a subject to self-ascribe  $c_1$  and  $c_2$  and to ascribe  $c_3$  to another subject. But this is a situation in which  $c_1$  and  $c_2$  are subject-unified and the pair,  $c_2$  and  $c_3$ , is not. This is not a situation in which the trio,  $c_1$ ,  $c_2$ , and  $c_3$ , is partially subject-unified. The same will be the case with respect to stream-unity. For any pair or trio of conscious states, they will either all be part of the same stream of consciousness or not. If they are, then they are stream-unified. If not, then not. There does not seem to be any way to understand what it would mean for them to be partially stream-unified. Thus,

the view that two mental mechanisms can each partially, but not completely, unify a subject's conscious states along any of these dimensions of conscious unity is problematic, because, in many cases, conscious unity seems to be an all or nothing proposition.

What about the alternative? What about the view that each of the two mental mechanisms is independently sufficient to produce any given kind of conscious unity? This view is also problematic because it amounts to the view that the two mental mechanisms causally overdetermine the different kinds of conscious unity? And this seems implausible in the face of the above considerations about the very different natures of the mechanisms. It also entails a rather significant resource cost, for it would entail that there are at least two independent mental systems doing the very same unifying work at any given time.

It is not exactly clear from Rosenthal's account how he conceives of the division of labour for the two mental mechanisms responsible for conscious unity. It is possible to read him as endorsing the view that they have overlapping explanatory domains, but for the above reasons, this view is problematic. An alternative view, according to which the two mechanisms have completely distinct explanatory domains, avoids these worries. Although it is possible to read Rosenthal as endorsing the overlapping view, nothing he says is directly inconsistent with the alternative view and so he would be free to accept it. The task would then be to provide more details about which kinds of conscious unity each mental mechanism is responsible for.

#### *Mere Conjunctive Co-Representation*

Part of what unifies conscious states, says Rosenthal, is the fact that *nots* often represent many lower-order states all at once. This invites a question about what these collective higher-order representations are like. Rosenthal has little to say about the precise nature of these higher-order representations. What I will argue here is that these collective higher-order representations must satisfy a certain condition.

Consider object-unity. The notion that object-unity might be the product of the co-representation mechanism has a high degree of plausibility. We know that our perceptual systems are modular in nature. Their representational contents at any given time will represent the particular physical properties they are respectively attuned to. What better to produce a unified experience of a single object from the disparate contents of these perceptual modules than a higher-order representation that incorporates them all into a single representation? It is important, however, that such collective higher-order representations are not merely conjunctive representations of the lower-order contents.

Suppose an object has a number of properties and that two of these properties,  $X$  and  $Y$ , are associated with distinct perceptual modules specialized to detect them. One way in

which the contents of these perceptual sub-systems could be co-represented by a HOT would be in a merely conjunctive manner. Such a HOT would have a content like the following: “I am in a mental state that represents that something is *X* and in a mental state that represents that something is *Y*.” This content is consistent with having a perception of a single object that is both *X* and *Y*, but it is also consistent with having a perception of two objects, one of which is *X* and the other *Y*. Merely appearing together in the same higher-order content is not enough to guarantee an experience in which *X* and *Y* inhere in the same object. What would seem to be required would be for the HOT to represent some kind of relation between *X* and *Y* or for the HOT to represent that they are instantiated by the same particular. The former could be achieved by associating the two properties to the same spatial location, e.g., “I am in a mental state that represents *X* as being at (location) *x* and in a mental state that represents *Y* as being at *x*.” The latter could be achieved by a HOT with a content like “I am in a mental state that represents (object) *a* as being *X* and in a mental state that represents *a* as being *Y*.” Both of these higher-order contents represent lower-order contents collectively, but they are more than mere conjunctive representations of the lower-order contents; there is a common element across the conjuncts to which perceived properties are associated. For our purposes, however, the details of how this association between *X* and *Y* is achieved is unimportant. What matters is that the higher-order representation do more than represent the two properties independently in the same content. It is also worth pointing out that a similar kind of cross-referencing would have to occur across mental states over time in order to successfully represent motion in addition to location.<sup>2</sup>

This requirement on the collective representations employed by the co-representation mechanism is not made explicit by Rosenthal. As has already been suggested, however, there is room for Rosenthal to accept it. There are a number of places where Rosenthal discusses the contents of HOTS and their target states, and in these discussions, Rosenthal allows for the possibility that HOTS do more than merely re-present the contents of their target states. The cocktail party effect and the case of the dental patient, both discussed earlier, are examples. Given that HOTS are free to add or subtract from the contents of their target states, there is no bar, in principle, to their representing certain relations as holding between lower-order states and their contents.

### *Mere Direct Reference*

The common-ascription mechanism unifies consciousness by exploiting the direct and indexical manner in which HOTS refer to the subjects of their target states: each HOT ascribes its target state to an individual via the (mental analogue of the) reflexive indexical “I” and so the subject of a HOT is left with the sense that its conscious states all have the same subject.

The impression one gets from Rosenthal’s account is that it is direct and unmediated reference *itself* that is somehow responsible for producing the sense of unity. Reflexive indexical reference may well have something to do with producing a sense of conscious unity, but indexical reference cannot be enough to do so by itself.

Consider a situation in which you hear a person utter the sentences “I am tall” and “I am the oldest in my family” in the same conversation. Given the context—you witness the same person perform the two speech acts—you naturally take the two tokens of “I” to refer to the same individual. However, it is not the mere fact that you witness two tokens of “I” in close temporal proximity to each other that you understand them to have the same referent. Suppose you overhear the same two sentences and that they are, in fact, uttered by the same individual, but you do not *know* that they are uttered by the same individual. Perhaps you are not in direct visual contact with the speaker and you do not recognize the voice as the same across utterances. In this case, the two occurrences of “I” would have the same referent, but you would not understand them in this way. What is important to notice in this situation is that both tokens of “I” refer to their referents in a direct and unmediated way and the way they get their semantic content remains fixed across occurrences, but this does not by itself determine that audiences will understand two separate tokens of “I” to have the same referent. What is required in addition is certain contextual information and a disposition on the part of the audience to apply a working knowledge of the semantics of “I” to that contextual information.

An analogous point can be made with respect to the mental analogue of “I.” Even though HOTS refer to the subjects of their target states in a direct and unmediated fashion via the mental “I,” nothing *necessitates* that downstream mental processes or larger mental systems take separate HOTS to refer to the same individual. Consider the mental demonstratives “this” and “that.” They refer to their referents in a direct and unmediated fashion, but this fact does not determine that minds in which they are tokened take separate tokens of these mental demonstratives to refer to the same object. In fact, it is most often the case that they are *not* taken in this way. If I spot a dark shadow in the distance while hiking in the forest I may think to myself “That could be a bear.” If a short time later I look at a cloud in the sky and think “That looks like a hat,” I do not take myself to have encountered a bear that looks like a fluffy white hat.

One might point out that there is a significant difference between the mental demonstratives “this” and “that” and the reflexive indexical “I.” Tokens of the latter always refer to the same thing when tokened in the same mind, whereas this is not the case with “this” and “that.” The content of the mental “I” is determined in a context-sensitive manner like

<sup>2</sup>Thanks to one of the reviewers for pointing this out.

“this” and “that,” but the contexts in which tokens of “I” occur ensure that all tokens that occur in the same mind always refer to the same object. Minds like ours need to accommodate varying references across different tokens of the mental demonstratives, but they do not need to accommodate variance across different tokens of the mental “I.” Thus, they treat the mental “I” differently than they do the mental demonstratives.

It is certainly the case that normal minds take successive tokens of “I” to refer the same individual. But this is not a *necessary* property of minds. Consider subjects suffering from dissociative identity disorder (otherwise known as multiple personality disorder). (See Brand and Loewenstein (2010), Kluft (1996), and Ross (1996) for general discussions of this disorder.) The memories, thoughts, decisions, etc., of these individuals are compartmentalized into distinct identities and personalities. Presumably these individuals often employ the mental “I” when they have thoughts about these mental states. Thus, subjects suffering from these disorders would seem to be examples of minds in which successive tokens of “I” are not taken by the same mind to refer to the same individual. What this shows is that it cannot be the mere fact that the mental “I” refers to its subject in a direct and unmediated way that is responsible for conscious unity. The mental systems that operate on hors have an equally important role to play.

I am not aware of any commitments Rosenthal might have that would prevent him from accommodating this point. Thus, neither this issue or the concerns I have raised above generate conclusive objections to his account. What they do, however, is reveal how his account should be developed and refined so as to avoid certain problems. I will now turn to some more serious worries that will reveal why Rosenthal’s account cannot, ultimately, be the final word on conscious unity.

### The Limitations of Co-Representation

As a means of accounting for conscious unity, the co-representation mechanism has a lot going for it. If a number of lower-order states appear in experience as unified in some way, it is plausible to think that their unification is the result of being represented together in a certain way. In the previous section, we saw how this might work in the case of object-unity. Similar considerations could apply with respect to spatial unity. However, there are other kinds of conscious unity for which the co-representation mechanism is either inadequate or for which it is not at all obvious how the mechanism could do the necessary unifying work.

Consider phenomenal unity again. My current global experience contains all of my current conscious states. Since all my current conscious states are part of the same global experience, they are all mutually phenomenally unified with each other. In order for the co-representation mechanism to

account for this, it would have to be the case that all of my current conscious states are represented together by a single higher-order state. But this seems highly unlikely. Our sensory experiences can be, and often are, highly complex. For a theory like Rosenthal’s, this would seem to suggest that, not only do conscious subjects like us have extremely complex first-order sensory states, but that all the content of those first-order states must be “mirrored” or accounted for via belief-like states at the higher-order level. And as was pointed out above with respect to mere conjunctive representation, the higher-order states must also explicitly represent certain spatial relationships between the various elements of experiences, both at a time and across time. This would place an extremely large computational burden upon any system that instantiates this mental architecture. And for what purpose? It is easy to see how having complex sensory states would benefit an organism in the kind of environment we find ourselves in, but what is the benefit to having *beliefs* or *thoughts* about all our sensory states?<sup>3</sup> Note, this is not a question about the benefit of having the *capacity* for thought—clearly the capacity for thought has great benefit—but a question about the benefit of constantly having all of our first-order states represented in thought. Neural tissue is metabolically expensive (Aiello & Wheeler, 1995). This makes it quite unlikely that systems embodying this kind of representational parallelism would evolve absent an evolutionary benefit. Carruthers (2000, p. 221) calls this “the objection from cognitive overload.”<sup>4</sup> Byrne (1997) presents an additional reason to doubt that we have the kind of complex higher-order thoughts this architecture requires. Many of our thoughts are unconscious, but we can, via introspection, make them conscious. Yet, when we try to introspect our thoughts about our sensory experiences, we never encounter such monstrously complex thoughts, especially not a single thought that represents *all* of one’s lower-order sensory states.

The shortcomings of the co-representation mechanism go even further. We all enjoy temporally extended episodes of consciousness, many of them last for hours at a time. All the conscious states in such an episode are stream-unified. In order for the co-representation mechanism to account for stream-unity, higher-order states would have to be capable of representing all of one’s conscious states in a given temporal span of consciousness at once. Even if, contrary to the above, there are higher-order states capable of representing all of a subject’s conscious states at a time, there would still be the question regarding a subject’s *past* conscious states. To the

<sup>3</sup>Seli (2012) makes a case for the utility of higher-order representations, but his focus is on the utility of having thoughts about thoughts, not about having thoughts about all our sensory states.

<sup>4</sup>It should be noted that Carruthers is himself a higher-order theorist. His own higher-order theory is developed, at least in part, as a response to this objection.



extent it is unlikely that a higher-order state can represent all of a subject's conscious lower-order states at a time, it is even more implausible that higher-order states are capable of representing all the conscious states in an entire episode of consciousness at once.

Rosenthal (2005c, p. 342) himself concedes the former. This is one of the reasons why he includes the common-ascription mechanism in his account, the idea being that it will pick up the slack where the co-representation mechanism falls short. This may seem to cover for the limitations of the co-representation mechanism, but as I will argue in the next section, the common-ascription mechanism is not up to the task.

### The Limitations of Common Ascription

Like the co-representation mechanism, the common-ascription mechanism has a lot going for it. An important part of our experiential lives and our sense of ourselves as persons is the fact that we view past experiences and events as important aspects of who we take ourselves to be now. We all have a "personal identity," to use a colloquial term, and we think of ourselves within the context of an extended narrative or personal history. This is certainly a kind of conscious unity—Flanagan (1992) calls it "strong self-consciousness"—and the way in which we refer to ourselves directly in thought would certainly seem to have something to do with it. The deeper problem for the common-ascription mechanism and for Rosenthal's account as a whole, however, is that it cannot make up for some of the apparent shortcomings of the co-representation mechanism. I will present two reasons for this.

#### *Expectations and Conscious Unity*

HOTS make their target states conscious, but HOTS are not generally conscious themselves. This poses a bit of a challenge for Rosenthal. How can unconscious states be responsible for conscious unity?<sup>5</sup> Rosenthal addresses this worry by distinguishing between an explicit and an implicit sense of unity. Subjects enjoy an explicit sense of unity when they actively introspect. It is then that some of their HOTS become conscious. This is associated with the explicit sense of unity, because when HOTS are themselves conscious, their contents, including the direct way in which they refer to the subjects of their target states, are also conscious. Most of the time, however, subjects of consciousness are not engaged in active introspection. But when they are not actively introspecting, maintains Rosenthal, they still enjoy a tacit or implicit sense of unity. The tacit sense of unity is what remains after one's HOTS are no longer conscious. But what, exactly, is this tacit sense of unity? It is, according to Rosenthal, the expectation we have that we can become actively aware of our mental states if we wish. Rosenthal (2005c, p. 345) sometimes refers to this as a "dispositional" sense of unity.

Because periods of active introspection are infrequent and short, HOTS can only be used as an explanatory tool for the kinds of conscious unity that are associated with periods of active introspection. Many other kinds of conscious unity, however, are much more persistent and pervasive throughout experience. This places a greater explanatory burden upon the tacit sense of unity. The question for Rosenthal is whether the *expectation* that one can, at any time, actively introspect one's mental states can do the explanatory work it needs to do.

As we have already seen, many of the different kinds of conscious unity have significant phenomenal consequences, and for these kinds of unity there is a significant phenomenal difference between a pair of conscious states being unified in the relevant way and not being unified in that way.<sup>6</sup> Some expectations also clearly have a phenomenal consequence. For instance, children seem to experience a certain set of feelings when they anticipate the opening of a jack-in-the-box. Likewise, certain feelings accompany the anticipation of a reprimand from a parent or superior. But other expectations have a much more subtle phenomenal consequence if they have a phenomenal consequence at all. When I step into an elevator and the door closes, I expect to feel a certain sensation in the pit of my stomach when the elevator begins to move. The sensation that accompanies the movement of the elevator certainly has an identifiable phenomenal character. The question, however, is whether the *expectation* I have beforehand has any sort of phenomenal consequence, and it is not at all obvious that it does.

There are others kinds of expectations that appear to be even more phenomenally inert. When I start my computer, for instance, I expect a certain sequence of events, but there is no *experience* of anticipating a computer's boot-up sequence. Or consider the expectations that I can add 2 and 3 together in my head if I decide to, that I can recall at will the births of my children, and that I can, if I wish, focus my attention on my big toe and wiggle it. None of these expectations seem to have any bearing at all on my current experience. Rosenthal associates the tacit sense of unity with the expectation that we can, at any moment, introspect our own mental states if we wish to. Many expectations do not seem to have any impact whatsoever upon the nature of our current experience. If they have no phenomenal consequence, it becomes unclear how a

<sup>5</sup>A similar kind of worry has also been raised for higher-order theories of consciousness more generally. How could unconscious states make us conscious of other mental states?

<sup>6</sup>Note that this does not entail that these kinds of unity have their own distinctive phenomenal character that they add to experience. Being unified in a certain way may effect the overall phenomenal character of an experience, but it can do so without adding a special "feel" of unity. One of the reasons for avoiding commitment to a special feel of unity is that it invites a kind of regress problem (see Hurley, 1998; Siewert, 2001; Tye, 2003).

mere expectation could be said to be responsible for many of the different kinds of conscious unity.

### *Developmental Considerations*

The second criticism I want to raise for the common-ascription mechanism is that it makes unity contingent upon two developmentally advanced mental abilities. As was alluded to above, a  $\text{HOR}$  is a thought with a content of the form “I am in such and such a mental state.” This means that, in order for a subject to have a  $\text{HOR}$ , the subject must (i) be capable of a certain kind of self-reference, (ii) have the representational capacity to represent the contents of its lower-order states in sufficient detail—we have already seen how this is a source of difficulty for the co-representation mechanism—and (iii) have the ability to characterize those contents as contents of a *mental state*. Self-reference and having the concept of a mental state are both relatively advanced mental abilities. The issue for Rosenthal is that some kinds of conscious unity seem to be present in experience even when some of the subjects of those experiences do not have the relevant mental abilities.

What it means for a subject to possess or have a concept is a matter of some dispute, but what is relatively uncontroversial is that having a given concept correlates with being able to correctly apply it in certain ways. With respect to the concept of a mental state, this would seem to require the ability to discriminate between one’s own mental states or to discern when one’s mental states differ from those of another. This, in turn, might require one to demonstrate certain levels of self-awareness, such as being able to tell when one is happy rather than sad. Or it might require the ability to discriminate between a current experience and a memory. Or it could require the ability to recognize that others have different perceptual perspectives and, as a result, may have different beliefs. In the developmental literature, this latter ability is associated with having a “theory of mind.” In a typical test for this ability, a young child and another individual are together shown the location of an object. The other individual is asked to leave the room and the object is moved to another location. The child is then asked where the person who left the room will look for the object when she returns. Children are only able answer this question correctly when they reach 3 to 4 years of age (Gopnik and Astington, 1988; Wimmer and Perner, 1983). It is unclear whether any other animals ever acquire the concept of a mental state. Some of the advanced social mammals, like primates and canines, who are able to respond to the displeasure of others in their social groups or cooperate in food-gathering and hunting activities, could be argued to have an extremely primitive concept of a mental state, but one does not have to descend too far down the mammalian hierarchy before almost all behaviour can be explained without reference to such a concept.

Some species of conscious unity seem to be present in

conscious experience long before subjects acquire the concept of a mental state or possess the ability to ascribe mental states to themselves. Two kinds of unity for which this consideration is particularly apt are phenomenal unity and stream-unity. With respect to the former, there is a strong case to be made for the claim that phenomenal unity is perfectly ubiquitous (Bayne, 2008, 2010; Friesen, 2013). To my knowledge, the only challenge to this claim comes from Lockwood’s (1989) interpretation of the split-brain data.<sup>7</sup> However, even those who challenge the ubiquity of phenomenal unity maintain that exceptions are quite rare, occurring only in atypical subjects like split-brain patients. This means that, for the vast majority of conscious subjects, including very young children and other mammals, phenomenal unity is an ever-present phenomenon. Thus, phenomenal unity cannot be explained by a mental capacity that very young children and many other mammals may not have.

The very same considerations apply for stream-unity. Conscious subjects, even very young children and many non-human animals, have temporally extended episodes of consciousness. Whatever is responsible for unifying conscious states into a stream of consciousness cannot depend on an ability they do not have. We saw that the co-representation mechanism cannot account for phenomenal unity or stream-unity. Thus it falls to the common-ascription mechanism to do the unifying work. But if this mechanism cannot account for the kinds of conscious unity that the co-representation mechanism is unable to account for, then there are some serious gaps in Rosenthal’s account.

Readers familiar with the literature on higher-order theories of consciousness will recognize the parallel between this objection and a well-known objection to Rosenthal’s  $\text{HOR}$  theory of consciousness (Siewert, 1998, section 6.5; Tye, 1995, p. 5; Seager, 2004). If conscious experience is a product of higher-order representation in the way that Rosenthal claims, then it follows that subjects incapable of  $\text{HORs}$  cannot be conscious. To many it seems absurd to claim that young children and most non-human mammals are incapable of conscious mental states. Gennaro (2004b) defends  $\text{HOR}$  theories by arguing that  $\text{HORs}$  do not require the kind of sophisticated cognitive machinery many believe they do. Carruthers (1998), on the other hand, bites the bullet and maintains that young children and animals do not have conscious mental states. To theorists like Carruthers, the parallel objection to Rosenthal’s account of conscious unity will have no sway—for creatures that do not have conscious states, questions about conscious unity do not arise—but for those who find the objection against  $\text{HOR}$  theories to have some force, the parallel

<sup>7</sup>Views according to which split-brain patients have two distinct streams of consciousness do not challenge the ubiquity of phenomenal unity. They are views according to which a single brain houses two subjects of experience, each of which enjoys a single phenomenally unified stream of consciousness.

objection to Rosenthal's account of unity will have just as much force.

### Representational Mismatches

In this final section, I will turn from a critique of Rosenthal's account as an account of conscious unity and make an observation about how his account of conscious unity is also relevant to the broader discussion of higher-order theories of consciousness. One of the main objections that has been raised against higher-order theories of consciousness is an objection I will call the "Representational Mismatch Objection" or, for convenience, the "Mismatch Objection." I will argue that Rosenthal's account of conscious unity relies upon the same kind of representational mismatches that the Mismatch Objection trades upon.

#### *The Mismatch Objection*

The Mismatch Objection has been pressed against higher-order theories by numerous theorists (see Byrne, 1997; Gois, 2010; Neander, 1998; Levine, 2001; Block, 2011). Whenever representations are involved, including mental representations, it is possible for the representational vehicle to misrepresent its target. Higher-order theories maintain that mental states are conscious when they are represented in the appropriate way. Given the possibility of misrepresentation, the question can be asked what it is like for the subject when a higher-order state misrepresents its target lower-order state. More specifically, the question can be asked whether, in cases of misrepresentation, the phenomenal character of the subject's experience conforms more to the representational content of the (higher-order) representing state or the (lower-order) represented state. No matter how it is answered, the answer seems to undermine the theory.

Suppose my visual system is working correctly while I look at my red mug under normal lighting conditions. This will generate a sensory state in me that represents my mug as red. Now suppose my sensory state becomes the representational object of a higher-order state, but, instead of representing my (lower-order) sensory state as a state that represents my mug as red, it represents my sensory state as a state that represents my mug as green. What will the phenomenal character of my experience be? Will it have the property of phenomenal redness or phenomenal green-ness? If it has the property of phenomenal redness, then it becomes unclear what the role of the higher-order state is in making the lower-order state conscious. For if the phenomenal character of my experience conforms to the representational content of my lower-order state, we are left to wonder whether I could have had the very same experience if the lower-order state had occurred in the absence of the higher-order state. The door is opened to the possibility that mental states do not need to be targeted by higher-order states to become conscious. Suppose, instead, that my experience has the property of

phenomenal green-ness, conforming to the representational content of my higher-order state. Since higher-order states are not, in general, themselves targeted by further higher-order states, the door is again opened to the possibility that a mental state can be conscious without being the target of any higher-order state.<sup>8</sup>

Many regard the Mismatch Objection as decisive. Others (Gennaro, 2004b; Rosenthal, 2011; Weisberg, 2011a, 2011b) have argued that it is not. I will not weigh in on this dispute here. Instead, I will show that Rosenthal's account of conscious unity systematically incorporates the kind of representational mismatches that the Mismatch Objection trades upon. I will do this by considering each of the two mental mechanisms in Rosenthal's account in turn.

#### *Conscious Unity and Representational Mismatches*

The co-representation mechanism purports to unify a series of conscious states by representing all of them together in a single higher-order representation. This scenario invites the Mismatch Objection because it is a situation in which the overall phenomenal character of a subject's experience at a time most closely matches the content of an unrepresented mental state. By offering an account of conscious unity, Rosenthal implies that conscious unity would not occur independently of the mechanisms featured in the account. In this case, it is implied that a situation in which a series of mental states are not represented together by a single HOT would generate a different kind of global experience than a situation in which those same states were represented by a single HOT. It is not enough to have the respective lower-order states each represented by their own distinct HOT. Rather, the unified character of the subject's experience is the result of the lower-order states being represented *together* by a single HOT. This constitutes a kind of representational mismatch. The representational content at the higher-order level includes a representational element that is not present at the lower-order level. Though all of the represented states appear at the lower-order level, their togetherness is not represented there. It is only represented at the higher-order level. Since HOTs are not themselves typically represented by further higher-order states, it would seem as though we have a situation in which the character of a subject's overall experience matches the representational content of an unrepresented state. And from this we are left to infer that it is the unrepresented higher-order state that is responsible for the character of the experience.

<sup>8</sup>Mandik (2009) has also raised an objection against higher-order theories that is driven by considerations of representational mismatches. Mandik's conclusion, however, is a bit stronger and perhaps a bit more general than the conclusions of those listed here. He argues that there is no such property as *being represented* and so there is no property for the property of *being conscious* or *being phenomenal* to be identified with.

The common-ascription mechanism also makes conscious unity a product of the content of higher-order states. According to this way of accounting for conscious unity, the sense of unity (explicit or tacit) hinges upon the fact that HOTs refer to subjects of their target states in a direct and unmediated way via the mental reflexive indexical “I.” Here again the phenomenal differences associated with conscious unity correspond to a representational element that occurs only at the higher-order level. The mental indexical this unifying mechanism hinges upon is a representational element in HOTs, not their target states. And just as with the co-representation mechanism, we are left to infer that it is the content of higher-order states that is responsible for the overall phenomenal character of experience.

It is doubtful that the observation I have made here about how Rosenthal’s account of conscious unity relies upon representational mismatches will sway many opinions about the force of the Mismatch Objection against his HOT theory of consciousness. Those who believe the Mismatch Objection to be decisive against it have merely been provided with more fodder for their cannon. Those who believe the Mismatch Objection is not decisive, will rely on the same arguments to show that these representational mismatches are not problematic. However, the observation I have made here does contribute something of significance to the debate. One defensive strategy the higher-order theorist could employ would be to argue that, although representational mismatches are possible, they are, nevertheless, exceptions to the rule. What my observation shows is that this strategy is off the table for the higher-order theorist who wishes to adopt Rosenthal’s account of conscious unity.

### Conclusion

Conscious unity is not a homogeneous mental property. There are many different kinds of conscious unity and any attempt to account for conscious unity must be sensitive to this. It is implausible to think that a proportionally small number of mental mechanisms will be able to do the job of accounting for all of them. A closer examination of the two mental mechanisms in Rosenthal’s account bears this out. There are at least some kinds of conscious unity that neither mechanism can account for. Thus, it fails as a perfectly general account of conscious unity.

### References

- Aiello, L. C., & Wheeler, P. (1995, April). The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution. *Current Anthropology*, 36(2), 199–221.
- Armstrong, D. M. (1968). *A materialist theory of mind*. London: Routledge.
- Armstrong, D. M. (1981). What is consciousness? In *The nature of mind and other essays* (pp. 55–67). Ithaca, New York: Cornell University Press.
- Bayne, T. (2008). The unity of consciousness and the split-brain syndrome. *The Journal of Philosophy*, 105(6), 277–300.
- Bayne, T. (2010). *The unity of consciousness*. Oxford: Oxford University Press.
- Bayne, T., & Chalmers, D. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration, dissociation* (pp. 23–58). Oxford: Oxford University Press.
- Block, N. (2011, July). The higher order approach to consciousness is defunct. *Analysis*, 71(3), 419–431.
- Brand, B., & Loewenstein, R. J. (2010). Dissociative disorders: An overview of assessment, phenomenology, and treatment. *Psychiatric Times*, 27(10), 62–69.
- Brown, R. (2014). The HOROR theory of phenomenal consciousness. *Philosophical Studies*, 1–12. doi: 10.1007/s11098-014-0388-7
- Byrne, A. (1997). Some like it HOT: Consciousness and higher-order thoughts. *Philosophical Studies*, 86, 103–129.
- Carruthers, P. (1996). *Language, thought, and consciousness: An essay in philosophical psychology*. Cambridge, England: Cambridge University Press.
- Carruthers, P. (1998). Natural theories of consciousness. *European Journal of Philosophy*, 6(2), 203–222.
- Carruthers, P. (2000). *Phenomenal consciousness: A naturalistic theory*. Cambridge, England: Cambridge University Press.
- Coslett, H. B., & Lie, G. (2008). Simultanagnosia: When a rose is not red. *Journal of Cognitive Neuroscience*, 20(1), 36–48.
- Dainton, B. (2000). *Stream of consciousness: Unity and continuity in conscious experience*. London: Routledge.
- Dennett, D. C. (1978). Where am I? In *Brainstorms: Philosophical essays on mind and psychology* (pp. 310–323). Montpelier, Vermont: Bradford Books.
- Flanagan, O. (1992). *Consciousness reconsidered*. Cambridge, Massachusetts: The MIT Press.
- Friesen, L. (2013). *The structure of consciousness* (Doctoral dissertation, University of Massachusetts, Amherst). Retrieved from [http://scholarworks.umass.edu/open\\_access\\_dissertations/794/](http://scholarworks.umass.edu/open_access_dissertations/794/)
- Gennaro, R. J. (Ed.). (2004a). *Higher-order theories of consciousness: An anthology*. Philadelphia: John Benjamins Publishing Company.
- Gennaro, R. J. (2004b). Higher-order thoughts, animal consciousness, and misrepresentation. In R. J. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology* (pp. 45–66). Philadelphia: John Benjamins

- Publishing Company.
- Gois, I. (2010). A dilemma for higher-order theories of consciousness. *Philosophia*, 38, 143–156.
- Gopnik, A., & Astington, J. W. (1988, February). Children's understanding of representational change and its relation to the understanding. *Child Development*, 59(1), 26–37.
- Hurley, S. L. (1998). *Consciousness in action*. Cambridge, Massachusetts: Harvard University Press.
- Kluft, R. P. (1996). Dissociative identity disorder. In L. K. Michelson & W. J. Ray (Eds.), *Handbook of dissociation* (pp. 337–366). Springer.
- Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Englewood, Colorado: Roberts and Company Publishers.
- Kriegel, U. (2009). *Subjective consciousness: A self-representational theory*. Oxford: Oxford University Press.
- Levine, J. (2001). *Purple haze: The puzzle of consciousness*. Oxford: Oxford University Press.
- Lockwood, M. (1989). *Mind, brain, and the quantum: The compound 'I'*. Cambridge, Massachusetts: Basil Blackwell.
- Lockwood, M. (1994). Issues of unity and objectivity. In C. Peacocke (Ed.), *Objectivity, simulation, and the unity of consciousness: Current issues in the philosophy of mind* (Vol. 83, pp. 89–95). Oxford: Oxford University Press.
- Lycan, W. (1987). *Consciousness*. Cambridge, Massachusetts: The MIT Press.
- Lycan, W. (1996). *Consciousness and experience*. Cambridge, Massachusetts: The MIT Press.
- Mandik, P. (2009). Beware of the unicorn: Consciousness as being represented and other things that don't exist. *Journal of Consciousness Studies*, 16(1), 5–36.
- Marks, C. E. (1981). *Commissurotomy, consciousness, and unity of mind*. Cambridge, Massachusetts: The MIT Press.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83, 435–450.
- Neander, K. (1998). The division of phenomenal labor: A problem for representational theories of consciousness. *Philosophical Perspectives*, 12, 411–434.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Puccetti, R. (1981). The case for mental duality: Evidence from split-brain data and other considerations. *The Behavioral and Brain Sciences*, 4, 93–123.
- Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, 49, 329–359.
- Rosenthal, D. M. (1997). A theory of consciousness. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 729–753). Cambridge, Massachusetts: The MIT Press.
- Rosenthal, D. M. (2000). Introspection and self-interpretation. *Philosophical Topics*, 28, 201–233.
- Rosenthal, D. M. (2003, June). Unity of consciousness and the self. *Proceedings of the Aristotelian Society*, 103(1), 325–352.
- Rosenthal, D. M. (2005a). *Consciousness and mind*. Oxford: Oxford University Press.
- Rosenthal, D. M. (2005b). Sensory qualities, consciousness, and perception. In *Consciousness and mind* (pp. 175–226). Oxford: Oxford University Press.
- Rosenthal, D. M. (2005c). Unity of consciousness and the self. In *Consciousness and mind* (pp. 339–363). Oxford: Oxford University Press.
- Rosenthal, D. M. (2011, July). Exaggerated reports: Reply to Block. *Analysis*, 71(3), 431–437.
- Roskies, A. L. (1999, September). The binding problem. *Neuron*, 24, 7–9.
- Ross, C. A. (1996). *Dissociative identity disorder: Diagnosis, clinical features, and treatment of multiple personality* (2nd edition ed.). New York: Wiley.
- Schechter, E. (2012, April). The switch-model of split-brain consciousness. *Philosophical Psychology*, 25(2), 203–226.
- Seager, W. (2004). A cold look at the HOT theory. In R. J. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology* (pp. 255–275). Philadelphia: John Benjamins Publishing Company.
- Seli, G. (2012). The utility of conscious thinking on higher-order theory. *Philosophical Explanations*, 15(3), 303–316.
- Siewert, C. (1998). *The significance of consciousness*. Princeton, New Jersey: Princeton University Press.
- Siewert, C. (2001). Self-knowledge and phenomenal unity. *Noûs*, 35(4), 542–568.
- Sperry, R. W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist*, 23, 723–733.
- Sperry, R. W. (1984). Consciousness, personal identity and the divided brain. *Neuropsychologia*, 22(6), 661–673.
- Treisman, A. (1998). Feature binding, attention, and object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353, 1295–1306.
- Treisman, A. (1999, September). Solutions to the binding problem: Progress through controversy and convergence. *Neuron*, 24, 105–110.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–141.
- Tye, M. (1995). *Ten problems of consciousness*. Cambridge, Massachusetts: The MIT Press.

- Tye, M. (2003). *Consciousness and persons*. Cambridge, Massachusetts: The MIT Press.
- Ward, R., Danziger, S., Owen, V., & Rafal, R. (2002, February). Deficits in spatial coding and feature binding following damage to spatiotopic maps in the human pulvinar. *Nature Neuroscience*, 5(2), 99–100. doi: 10.1038/nn794
- Weisberg, J. (2011a, July). Abusing the notion of what-it's-like-ness: A response to Block. *Analysis*, 71(3), 438–443.
- Weisberg, J. (2011b). Misrepresenting consciousness. *Philosophical Studies*, 154, 409–433.
- Wimmer, H., & Perner, J. (1983, January). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128.