Representation, Consciousness, and Time

I. INTRODUCTION

The relationship between consciousness and intentionality continues to draw the attention of philosophers. One interesting and unusual approach is offered by Bourget in *Noûs* (2010, 44:1, pp.32-58) who argues that all possible conscious states are underived if intentional. This is a component of what he calls the *Originality thesis*, and, ultimately, his view that consciousness is “nothing more than” (p.33) a Pure Underived Representational Event (PURE). An underived state is one of which it is not the case that it must be realized, at least in part, by intentional states distinct from itself. In other words, an underived intentional state is one whose content does not metaphysically depend on any other intentional states (2010, p.33). Bourget gives both intuitive and empirical arguments for the Originality thesis, and here I offer skeptical replies. Despite the ingenuity of these arguments, they do not succeed in showing that conscious states are underived. I will argue that a plausible rejoinder to the PURE theory pivots on the claim that phenomenal consciousness, by necessity, has a temporal structure. Although my appraisal is negative, an examination of the PURE theory will throw new light on broader issues, including the unity of consciousness and the doctrine known as “phenomenal atomism.”

II. PURE REPRESENTATION & THE CONCEPTUAL ARGUMENT

Clearly, not all representations are conscious, e.g. words and pictures, and the same can presumably be said for many representational states of the mind and the nervous system generally. So, what makes a representation a phenomenal state? Bourget suggests it is when its intentional content is “original or underived…phenomenal states are simply underived states” (2010, p.33). The PURE theory thus consists of the following two claims:

1. *The originality thesis*: all possible phenomenal states are underived intentional states.
2. *The derivation thesis*: all possible non-phenomenal intentional states are derived.

Since the Derivation thesis is equivalent to saying that all underived intentional states are conscious, its conjunction with the first claim entails that conscious states are one and the same as underived intentional states. By “derived” Bourget means that the state’s intentional content is not intrinsic, which is to say that it depends on, or obtains in virtue of, the content of other representational states, e.g. as in how the representational content of ink on a page is parasitic on the intentionality of the author’s thoughts (2010, p.33). The PURE theory is compatible with there being non-conscious mental states, however, these would have to be derived in some sense.[[1]](#footnote-1)

Bourget further explains by surveying four ways intentional states could be derived (he does not say these are exhaustive). These are: (a) *Deferential derivation*. This is when mental contents depend upon the mental states of others, such as linguistic conventions, (Burge 1979, Putnam 1975); (b) *Compositional derivation*. Some representations are concatenations of intentional components (as in composite beliefs, such as “There’s a cat and a hippy on the mat,” or mental states that combine beliefs and experiences, as in the state of believing-that-P-while-experiencing-red; (c) *Dispositional derivation*. Intentional states can also be causally related to other intentional states. For example, perhaps nonoccurrent beliefs are (at least partly) dispositions to token certain occurrent beliefs. In this case, the intentionality of nonoccurrent beliefs would be dispositionally derived; (d) *Matching derivation*. Given the distinction between wide and narrow mental content, perhaps the former are derived from the latter in the sense that what determines a wide content is its “descriptive matching” to a narrow content. For instance, perhaps the content of the term “water” has the wide content H20 in virtue of the fact that H20 descriptively matches the narrow content about a colorless, odorless, potable liquid (2010, p.35). In what follows I will argue that phenomenal states must be derived, most likely in either sense (b) or (c).

III. CRITIQUE OF THE CONCEPTUAL ARGUMENT

The main argument for the originality thesis is straightforward: given a certain metaphysical intuition, namely that “any sensory experience could in principle occur in isolation from all other intentional states” (2010, p.36), it follows conscious intentional states can be metaphysically isolated, that is, it is not necessary that any sensory experience be realized by, or depend on, other intentional states. It would then follow that, necessarily, if a state is a conscious representation, then it is underived.

The intuition is supported by a thought experiment about a conscious subject who appears out of nothing, and then disappears after “a split second” (I hereby dub this being “Blinky”). According to Bourget (2010, p.36) it is possible that Blinky has only one representation, in this case, momentary visual consciousness of a white surface:

Take for example my current experience of the whiteness of my computer’s screen (call it *e*). On the face of it, a being could pop into existence, have a token of *e* for a split second, then disappear without having had any other intentional states or stood in any nontrivial relations to other intentional states.

If this is cogent, then it is possible Blinky’s experiential state could exist without metaphysically depending on, or being realized by, any other intentional states or relations to other such states (he only has one, after all). It would seem to follow that Blinky’s experiential state is underived. The point is assumed to generalize such that all other forms of sensory and non-sensory consciousness could also exist in isolation. However, a problem with the key thought-experiment is that it is underdescribed in a way that, I will argue, greatly diminishes its plausibility.

Bourget characterizes the PURE theory in terms of states rather than processes: only *states* are underived. The contrast between states and processes is between an instantaneous “snapshot,” or time-slice, lacking any temporal duration, and a series of metaphysical snapshots which, taken together, that is, linked by certain dependency relations (such as causality or constitution), comprise a temporally extended event. For example, he mentions that a mental process, such as thinking or reasoning, has a “temporal structure,” connecting the states that realize it. He is also ambivalent as to whether we should count a process as a kind of state, though it doesn’t really matter, since “If it is counted as a (temporally structured) state, it is a derived one, because a reasoning process is of necessity realized by the states it connects.” On the other hand, if processes are not states, then they have no bearing on the premises and theses of the PURE theory.

While I find this helpful and persuasive, my contention is that the central thought experiment trades on ambiguity between states and processes in the phrase “split second.” On one interpretation, a split second is literally an instantaneous moment. But this may not be what Bourget has in mind. Perhaps he just means the briefest possible duration for a conscious experience to occur. In turn, there is the question as to whether this is to be understood in terms of biological possibility or metaphysical possibility. Presumably it is the latter. But reflecting for a moment on the biological limitations will help clarify what is at issue.

Recent estimates of the shortest interval in which a subject can notice a conscious difference is in step with James’ report of 44 ms over a hundred years ago (1890: 613) and put the range between 25-150 ms (Coren, Ward and Enns, 2004: 351). The details need not distract us, but at some point, stimuli presented too briefly are not consciously apprehended. Of course it is reasonable to suppose that this limitation is only a contingent fact about the human nervous system. If something as far-fetched as Blinky is conceivable, then it is a small matter to additionally suppose the biologically required interval could be compressed. Now, could the interval be compressed so completely that we can conceive of Blinky’s conscious state as a durationless instant? I contend it is highly implausible that a conscious state could be (metaphysically, literally) instantaneous. After presenting my reasons I will then explain why this means conscious states cannot be isolated. My counterargument goes like this:

1. Experiences necessarily take some amount of time

2. Given (1), the Originality thesis is false.

So, 3. The Originality thesis is false.

Next, I will develop the case for the premises.

Bourget contends that his metaphysical intuition about Blinky is clear and convincing. If I can show that this is not the case, then the burden of argument shifts back to him. One basis for skepticism begins with the thought that consciousness presupposes dynamics: like a tennis volley, it is process-like, rather than state-like. This is because an instantaneous Blinky would be like a person frozen in time. To motivate this thought, suppose God hit the universe’s “pause button” and time ceased to move forward. One consequence is that all causal processes, including the neurophysiological basis for consciousness, would be stopped in their tracks. Would there still be something that it is like for us? I doubt it. If “frozen” while listening to some music, it seems inappropriate to say that I would continue to hear and enjoy the song even though I am stuck there, inert. Perhaps it could be argued that if I am paused while hearing a certain note, I would continue to be conscious of that audible content until I am “thawed.” But it seems to me that there is a difference between the pause-scenario and a case where I continue to experience the same content over and over. To be consciously aware of a certain note played continuously is just not the same thing.

Temporal extension is plainly a key feature of our phenomenology and suggests there is nothing that it is like to experience a literal instant. Resistance to this claim is perhaps attributable to a part-whole fallacy: just because a series of instantaneous events is conscious, it does not follow that the same is true of each instant, considered in isolation. In any case, I don’t need to show that this *must* be the case in order to challenge the Originality thesis. It is sufficient to merely show that the coherence of an instantaneous Blinky is dubious. Of course, almost any theorizing about time and consciousness is contentious and some would deny that there is necessarily any temporal structure. Although I find the intuition I am emphasizing very persuasive, perhaps there would be something that it is like within a durationless moment after all. However it suffices that there are other metaphysical intuitions about consciousness and time that seem to clash with the one about Blinky. So then it is not clear why we must accept the one rather than the other. At the very least, we don’t know what this might be like, but then I don’t see why we must rule out the possibility that it isn’t like anything.

Here is another example. Suppose two people are eating spicy food and are experiencing discomfort as a result. This time God only freezes time for one of the diners. In terms of avoiding discomfort, which is better, to be the one who is frozen or unfrozen? It seems to me clearly to be the former. Similarly, suppose a loved one is experiencing great suffering and your sole course of action is to activate a marvelous technology that somehow stops time, but only for that person. Have you helped them? It would seem you have—certainly and without hesitation I would want to be frozen if I was the loved one, assuming that nothing else could be done. These points make sense under the assumption that it is absurd to attribute consciousness to instantaneous states.[[2]](#footnote-2)

This view is in keeping with the idea that consciousness has an ineliminable temporal character, as in James’ “specious” or “practically cognized” present that has a “short duration of which we are immediately and incessantly sensible”(1890: 631). For James, and many others, consciousness is “no knife-edge, but a saddle-back, with a certain breadth of its own…” (1890: 609-10).[[3]](#footnote-3) Diverse philosophers (perhaps most notably James and Husserl, see Anderson and Grush 2009) have sought to explain the seeming awareness we have in consciousness of succession, persistence, and change. One’s experiences seem to be packaged in temporal intervals. Experience is somehow both past and present at once—not just as a contingent biological fact, but as a metaphysical necessity. This establishes the first premise in my argument.

However, we should proceed with care. The minimum duration of an experience could be understood in terms of either real time or represented time, meaning the first premise can be interpreted in two ways. This might make a difference to the overall argument. In light of the representational vehicle/content distinction, consider first how the felt duration of an experience might be due to the duration of its vehicles. Alternatively, perhaps felt duration is due to represented duration. On the second option the vehicle of experience is durationless but has content representing a temporally extended event. These would seem to be the prime candidates when it comes to how time is represented in consciousness. But I will argue that on either view any conscious state will have contents that are compositionally derived.[[4]](#footnote-4) With this distinction in mind I will now consider the justification for premise two.

The first version of the argument begins with the premise that the vehicles of conscious experience necessarily span some time. For example, perhaps Blinky’s consciously seeing the patch of white consists in tokening a series of instantaneous white-representing states at t1, t2, and so on for remainder of the felt interval. The second premise says that, given this first premise, the Originality thesis is false. This would seem to follow since the conscious content would be compositionally derived. But why exactly are temporal dynamics unfriendly to the Originality thesis?

Representing white throughout a conscious episode would require there to be distinct vehicle tokens of the same type. But this means Blinky cannot undergo a solitary conscious state with a single token content. The content of the whole experience would metaphysically depend on multiple and successive time slices, each of which individually represents white. An individual time slice is not itself a conscious experience of anything. It is only the series that is conscious. This is why temporal extendedness matters: the content of the conscious experience depends on the series of representational properties corresponding to each slice in the series of tokens.[[5]](#footnote-5) Bringing this back to the Blinky scenario, the experience of seeing a patch of white, however briefly, would seem to imply an interval from a past moment to a present moment and distinct token states and contents corresponding to each moment. But this means that the content of the conscious experience is derived from a succession of states that represent whiteness. Hence the second premise is established: if experiences necessarily take some time, then their contents are derived and the originality thesis is false.[[6]](#footnote-6)

A similar rationale holds for the other interpretation of the felt duration of consciousness. Perhaps instead of a succession of vehicles there is a single (instantaneous) vehicle representing the passage of time. Once again if Blinky’s visual consciousness always has a temporal component, then its content always has a complex structure and this would also be a case of compositional derivation. In addition, the representation of the passage of time might itself be derived, e.g. for there to be an experience of change (or persistence) both past and present must somehow be represented. Here too then Blinky’s experience of whiteness is realized by “other intentional states or relations to other intentional states” (2010, p.34) and thus has complex content by Bourget’s own lights.

I have argued that if consciousness has a temporal structure this puts pressure on the claim that its contents can be metaphysically isolated. The key step in the counterargument is that this commits us to intentional states with compositional contents. Bourget even appears to agree that a temporally structured state is necessarily a derived state (2010, p.41). This seems like a very strong case against the Originality thesis, however, next consider a possible reply. In a footnote (n.14, p.55), Bourget anticipates an objection that may seem promising. He contends that “total” experiences are not compositionally derived, since they are only determinates of their determinables. If this is correct, then perhaps even temporal structure can be understood in a way that does not imply that phenomenal states are derived.

Recall that the determinate/determinable distinction is a type of relation between the more specific and the less specific. However, unlike species/genus, the determinate is not a composite of, or “realized by,” its determinables (e.g. though “bachelor” is realized by “male” and “unmarried,” “color” does not realize, and is not a composite of “red.” Another example is shape, which is the determinable for the determinate round. Bourget (2010, fn.14, pp.55-6) suggests that “total experiences [e.g. the overall visual scene of a cluttered office] are determinates of their (alleged) components (their determinables).” So the total experience of a certain red square would be the determinate of its determinables, namely, the experience of redness and squareness, without being a composite of, or being realized by those experiences. This is why Bourget denies that total experiences are realized by component experiences. (If literally every phenomenal state is underived, then total experiences must not have components.) Hence, Blinky’s total experience would merely be a determinate of either experiencing whiteness, again and again (as in the first reading of premise one), or, experiencing whiteness and experiencing the passage of time (as in the second reading). Bourget also remarks that if determinables do sometimes realize their determinates, this would have to be a way that differs from other kinds of realization (p.56)

However, I find this reasoning unpersuasive. It is compatible with something standing in a determinate/determinable relation that it is a composite of something else in a sense relevant to the derivation of intentional states. The experience of a certain red square could still be a composite of an experience of a certain shade of red and the experience of a certain square shape even if it is not a composite of the experiences of redness and squareness. The example he gives of a compositional state was experiencing-redness-while-believing-that-P. That state has as a component a state that “has intentional directedness at least towards P” (2010, p.34). It seems to me that the simultaneous representation of shape and color also suggests compositional content. Since the color content could be changed without any change in the shape content, these must be composite, in the same sense of specific directedness: we individuate intentional states by their contents. Hence, experiencing redness is directedness towards red; experiencing squareness is directedness towards squareness, and so the state of experiencing-redness-while-experiencing-squareness is derived from these intentional components. This makes sense. After all, if a belief and an experience can form a composite, and a belief and a belief can form a composite, then why not two experiences?

An additional difficulty with Bourget’s presumed response is that my criticism is not that total experiences are composites of component *experiences*, but rather that the smallest “unit” of experience is a composite of *intentional states* that are not themselves phenomenal experiences. The determinable of a determinate experience must also be an experience, but the components of experience are not necessarily experiences. The argument for PURE representation needs to show that experiences are not realized by component representational states, whether or not those components are experiences. But the determinate/determinable distinction does not bear on that question. So, the issue is just beside the point of my counterargument.

 Perhaps it would make more sense for the PURE theorist to drop this line of reasoning and embrace phenomenal atomism—that is, the view that the phenomenal field can be decomposed into states that are independently conscious. Atomism is compatible with underived intentionality, but rejects the Originality thesis since their combinations are not assumed to be independently conscious—the total phenomenal field is a composite of the “atoms” (Bayne 2010, pp.225ff.). Meanwhile, the Originality thesis makes a very strong claim about all possible phenomenal states (i.e. including total experiences). If individual phenomenal states can occur in isolation, then it is natural to suppose nothing needs to be added for their combination to be like something.

Yet even this weakened version of the Originality thesis continues to face the counterargument about felt time; any putative atoms would not really be metaphysically independent since they would have compositional contents owing to either vehicle duration or represented duration. Either way, Blinky’s contents are not attributable to a unitary underived state and we can’t really conceive of fully isolated conscious contents.

IV. FURTHER CONSIDERATIONS BEARING ON THE ORIGINALITY THESIS

Besides the argument about Blinky, Bourget also provides some everyday examples of non-sensory experiences including tip-of-the-tongue feeling (TOT), experiencing familiarity, and conscious awareness of causation, e.g. that “it seems possible to be phenomenally conscious merely as of *something causing something* without undergoing other experiences or having any other intentional states” (2010, p.36). These are all arguably better described as cases of compositional derivability. For example, Gray (2002) compares the TOT phenomenon to a blindspot in the visual field, in that the subject has an expectation that something should be there, but isn’t. Gray explains this in terms of a failed attempt at intermediate-level processing, where a top-down template fails to find a bottom-up perceptual match. Representing the contrast between what ought to be there, and what is (or isn’t) there, implies compositional content since there are distinct representational states with their own intentional objects (Bourget 2010, p.34). Hence, the state of *expecting-that-P* is distinguished from the state of *representing-that*-*not-P* thus making TOT composite. Perhaps it is also a case of causal or dispositional derivation insofar as the representation of *not-P* is generated by the expectation preceding it.

Meanwhile, the experience of one thing causing another is also case of either compositional, causal or dispositional derivation, since it depends on representations of distinct events, and then a further representation of one bringing about the other. Consider the phenomenological difference between undergoing the ventriloquist’s illusion, where the voice is experienced as coming from the dummy’s mouth, versus merely seeing the dummy move while hearing somebody’s voice. Again, since the content of hearing the voice *as coming from the dummy* can be added or subtracted without altering the other contents, it would seem that this representation is a composite and causally dependent on other representations.

V. MINIMAL FUNCTIONALISM AND PURE CONSCIOUSNESS

Bourget also provides an argument for the Derivation thesis, that is, that non-phenomenal intentional states (i.e. attitudes) are always derived or parasitic on other mental states. Though not intended to directly argue for the Originality thesis, there are claims and inferences in it germane to that issue.

The argument for the Derivation thesis seems to go like this (2010, pp.41-2):

1. You (the reader) can experience the bent look of a stick immersed in water without tokening any attitudes about its being bent (especially beliefs and desires).

2. You (the reader) can experience mental imagery of a beach, without tokening any attitudes about any beach.

3. Meanwhile, to token any attitude (such as the belief that the stick is bent) conceptually requires certain dispositions to token other attitudes or phenomenal states.

4. So, the absence of attitudes about a given content is compatible with phenomenal experience of that content (by 1, 2 and 3).

5. So, “Attitudes are parasitic on the states they connect” (p.42) (by 3).

And, (implicitly),

6. So, Phenomenal states are not parasitic on the states they connect (assuming the point of 1 & 2 can be generalized to cover all experience, and, premise 4).

The upshot is that one can have an experience without tokening other intentional states, this is not the case when it comes to attitudes and their contents.[[7]](#footnote-7)

This is a simplified presentation tailored for a specific purpose (e.g. Bourget gives additional reasons for statement 3, and ultimately the central conclusion that attitudes are parasitic). My interest is that the examples about seeing the stick and imagining the beach involve claims about phenomenal experience that bear directly on the plausibility of the Originality thesis. Accordingly, this section shall focus on steps 1 and 2, and the inferences to 4 and 6.

 These examples seem to be asking us to consider the case of a more or less standard human being (“suppose that you are lying in bed…” 2010, p.41), so I will proceed under that assumption. One point is that Statements 1 and 2 only refer to *actual* states, not *dispositional* states. Even assuming that I can see a stick as bent, or imagine a beach, without actually tokening any attitudes, it does not follow that I have absolutely no *dispositions* to token any attitudes about those contents. Additionally, it is questionable whether we can experience sticks and imagine beaches without tokening attitudes. In fact, there probably would have to be actual distinct mental states, as well as dispositions to token other mental states. For instance, in considering the example about the beach, it is reasonable to assume that certain details are part of the imagery: Is it of a beach by day or by night? Which season is it? Are there are clouds in the sky? The point is that conscious imagery has contents that seem to demand the tokening of attitudes, or at least dispositions to token attitudes, such as the belief that the image is of a beach by day, or the disposition (and intention) to answer “day” when asked about it, and so on.[[8]](#footnote-8) In general, it seems implausible that I could have experiences or mental images concerning which I had no attitudes whatsoever. How can I fail to have the thought that it was a beach I was imagining, for instance, and not something else, or, at least dispositions to form such thoughts? This suggests that phenomenal states may be entangled in relations to other mental states much as attitudes are. If so then these examples fail to support the distinction between attitudes and phenomenal states. This is why I find that Bourget’s discussion of the Derivation thesis has the unintended result of further weakening the basis for the Originality thesis.

VI. PURE REPRESENTATION & THE EMPIRICAL ARGUMENT

Bourget also appeals to certain empirical considerations for additional support of the Originality thesis, such as a comparison between humans and other animals (2010, pp.36-7). Although our perceptual systems are similar to those of very simple-minded creatures, human cognitive systems are quite unusual and elaborate. Bourget takes this to suggest that phenomenal experience “can occur without thoughts.” (p.36). Another line of argument draws on the modularity of perception.[[9]](#footnote-9) Thus, blindness has little to no impact on the other senses, and even within a sensory modality deficits are selective (e.g. color can be subtracted from visual experience without affecting anything else). Similarly, cases of perceptual illusion also suggest sensory experience is “largely independent of influence from background beliefs” (p.36). Tying these points together, Bourget argues as follows:

The fact that sensory experiences are by and large nomically dissociable from one another suggests that one could in principle have any sensory experience in isolation from all other possible sensory experiences. Given that sensory experiences are also independent from thoughts, it is tempting to conclude that any sensory experience could in principle occur in isolation from all other possible intentional states. It would then follow that sensory experience is underived (p.37).

In other words, there are many examples of “nomically dissociable” sensory experiences, and this warrants (inductively) the inference that all possible phenomenal experiences are metaphysically independent.

 A general difficulty with this argument is that these are not unambiguous examples of nomically dissociable sensory experiences. Showing that the remainder of one’s total conscious state is largely unaffected by the loss of one (or several) subsystems is not the same as showing that the missing subsystem could support phenomenal consciousness all on its own. I will explain what I mean in considering replies to the specific cases he utilizes and argue that his empirical argument begs the question.

 Start with the comparison to animals. It is far from obvious that simple-minded organisms are phenomenally conscious. Animals more similar to us perceptually also seem to be more like us when it comes to their cognitive systems; similarities come in degrees after all. For example, although other mammals and birds have binocular color vision, human vision exhibits the architecture and other qualities specifically characteristic of primates (Van Essen and Deyoe, 1995). Since mammals and primates are most like us cognitively, perhaps it is reasonable to draw a different conclusion: might conscious experience depend on having processing that is at least somewhat like human beings? This seems likely from the example of blindsight and other implicit mental processes. Fairly sophisticated perceptual mechanisms can function in the absence of phenomenal consciousness, which in humans seems closely tied to reasoning, action, and speech. Although this might be mistaken—perhaps so-called implicit processes *are* conscious but not introspectible—skepticism is prima facie plausible when it comes to consciousness in organisms with systems that are *much* more rudimentary than the residual and implicit processes that seem to be unconscious for humans. As it seems plausible many simple minded animals are not conscious, this does not support the view that sensory systems are isolatable.

It is also not the case that the Originality thesis is supported by the fact that perceptual illusion is somewhat encapsulated from cognitive processes. That perception is only *largely* impenetrable suggests cognition has *some* influence, as with the fact that cultural programming affects whether one is susceptible to the Müller-Lyer illusion (Robbins, 2009). The Originality thesis requires a much stronger claim, namely that perceptual illusion could occur in total isolation. Cases of partial encapsulation do not establish that since they are compatible with something having merely derived intentionality. Derived states can also be partially encapsulated.

 As for the modularity of perception it is consistent with two quite different theses concerning the dependence of phenomenal states on other intentional states. Perhaps sometimes a part *p* of a system S can be F *whether or not* it is integrated in S—by “integrated” I mean *p* bears the right kind of metaphysical relations to other parts of S (such as causality or constitution); call this the *Independence* *principle*. There is also the *Dependence principle*, which states that a part *p* being F depends on its being integrated within a system S. The question is: when it comes to phenomenal consciousness, do instances of modularity support the Independence principle or the Dependence principle? According to the Independence principle, consciousness is analogous to the illumination of a bulb in a string of Xmas lights. Any burnt bulb is encapsulated in the sense that it has no effect on the ability of the others to light up. In addition, any bulb could be illuminated independently of the network (assuming it has its own power source). On the other hand, we also have the Dependence principle which is illustrated by the value of a quarter: Although removing a quarter from a monetary system leaves a quarter-sized deficit in the system’s total extant value, the quarter has no value independently; its economic value is clearly derived. Herein lies the problem: any deficit in the phenomenal field is compatible with, and expected on, either the Dependence and Independence principles. Hence, phenomenal field deficits license no inference either way. To claim, for instance, that modularity supports Independence is as unfounded as saying that something’s being a quarter is metaphysically independent just because removing it from an economy leaves a 25-cent hole. [[10]](#footnote-10) In other words, that removing *p* from S entails that S is not F, does not raise the probability (or otherwise support the claim) that *p* can be F independently of S. Bourget’s argument begs the question by assuming otherwise.

The modularity of perception suggests *the perceptual system* can mostly continue to do its job absent one (or several) of its subsystems.[[11]](#footnote-11) But this does not support the claim that its parts and subsystems could continue to produce experiences when isolated. To assume otherwise commits a fallacy of division. By analogy: although a carburetor could continue to be used to mix air and fuel after its removal from a Cessna 172, carburetors do not fly on their own, not even a little bit. Another example is the bilge pump on a boat—it doesn’t float, even a little, though its absence could cause the boat’s ability to float to gracefully decay—for a while.

Similarly, consciousness may be a property of the overall system, not its various parts. The analogy to money again illustrates. A quarter can be subtracted from an economy, but a quarter cannot exist in metaphysical isolation (a round piece of metal can be isolated, but a *quarter* derives its economic value from its functional role in an economy). In short, just because a functional state can be subtracted from the modular system of which it is a part, doesn’t imply that it can exist, qua functional state, in isolation from that system. Likewise, the intentional content had by a quarter—i.e. that it represents .25 cents, depends on how it’s positioned in a monetary system. Again, if the contribution of perceptual subsystems to consciousness is like a carburetor facilitating flight, or a quarter helping to sustain an economy, then isolation is not an option.

In reply, Bourget might claim that these analogies miss their mark because the functionality and intentionality is merely derived (e.g. 2010, p.52). In an example about the isolated early visual system of a mouse, he finds that “it would seem farther from having intentional states than today’s digital cameras.” The reason is because derived intentional systems need to be “integrated within larger perceptual and cognitive systems” (p.52). He makes a similar point about blindsight and other implicit processes—either they are not really mental, or not really intentional, or perhaps they are conscious after all. Yet the first two suggestions seem to be non-starters. Assuming the earlier, more rudimentary, subystems or residual islands of cortical neurons responsible for e.g. visuo-motor response are derived, why would the neurologically more complex later states be underived? Early systems for e.g. detecting edges or corners are aren’t tied into networks of dispositions in the way that attitudes or phenomenal states are, but they make a conscious difference when dissociated. On what basis are we to ascribe consciousness to one tangle of neurons, but not another (such as the *later* part of a mouse’s visual system responsible for conscious motion detection)? Perhaps it makes more sense to bite the bullet and insist all the little subsystems are all conscious after all. In this way, conscious intentionality could still be nonderived.

However, to assume there must be something that it is like to be a visual system fragment would beg the question in favor of the Originality thesis. Perhaps instead intentional states divorced from perceptual systems have no roles or contents (whether derived or nonderived). But my criticism doesn’t depend on claiming these system fragments are mental, underived or conscious. The point is that empirical considerations about the modularity of perception do not support the Independence principle. It is only if we assume that subsystems can be conscious on their own that their dissociability can be used as evidence for metaphysical independence.

VII. IMPLICATIONS FOR THE UNITY OF CONSCIOUSNESS AND ATOMISM

Conscious atomism is the view that the phenomenal field can be decomposed into units that are independently conscious (Bayne, 2010). Although Bourget’s account is not atomistic, his Blinky thought experiment is essentially friendly to atomism. The Modal argument for atomism turns on the apriori consideration that any phenomenal experience part of a subject’s total phenomenal state can be metaphysically isolated from the rest. That is, we can always conceive of subtracting any individual phenomenal state from a subject’s total experience without changing anything else. Bayne (2010, p.243) considers the example of a pain in one’s left leg—the atomist contends that one’s total experience could be otherwise exactly as it is minus that one pain. This might seem to be best explained by the fact that the pain, as an atom of consciousness, could, in principle, exist independently. However, Bayne finds this suggestion dubious, since the “pain in one’s left leg can occur only in the context of certain other experiences (such as an overall sense of one’s body).” This, however, is an unsatisfying response for three reasons. First, as an appeal to brute intuition, it is inconclusive—what of atomists who differ? Second, indeed it seems plausible that a pain’s having a certain character, intensity, and duration, could occur independently of one’s bodily awareness. It is not hard to accept that most of one’s body could be numb, or that there could be a peculiar disorder of proprioception and the kinesthetic sense. Third, Bayne later observes that various pathologies indicate phenomenal unity doesn’t depend on a sense of embodiment (2010: 268). Since the experience of one’s body as an integrated whole is not necessary, I don’t see why pains and other bodily feelings that normally have specific locations couldn’t be conceived to occur in isolation. This stands to reason given that the body can be experienced as fragmented.

Despite these worries I share Bayne’s skepticism about the Modal argument. Instead I suggest we recycle my response to the PURE theory: consciousness always has a temporal structure, and therefore the minimally conceivable units do not have atomistic representational contents. Having said that, something close to atomism might still be true. Though this is not the place for an examination of the Modal argument, we could consider whether it might show there are “temporal atoms” composing conscious subjects.

VIII. CONCLUSION

I have argued that Bourget’s central thought experiment is underdescribed when it comes to the temporal structure of experience and the nature of Blinky. In filling in these details, it is evident that the desired metaphysical intuition is either absent, greatly weakened, or conflicting with other, stronger, intuitions. Either Blinky’s experience is not isolated, or, if it is isolated, it is not clearly an experience. I have also tried to show that, despite Bourget’s other examples and arguments, it is rational to withhold agreement with the Originality thesis. Is this to say that consciousness is derived? This would be a hasty conclusion to draw.

Although I am highly critical of Bourget’s case for the PURE-ity of consciousness, I haven’t disproven it either. Consciousness is strange and diverse views ought to receive a hearing. We should be cautious about rejecting suggestions as to how we can improve understanding even if they seem bizarre. The PURE theory is certainly interesting enough to provoke a response and stimulate further discussion about broader issues, such as the unity of consciousness.

Still, I am doubtful as to whether the isolation framework can be trusted to yield the target intuitions. This is because there is even a sense in which derived intentionality can be isolated. To adopt an example of Dennett’s (1988), the very first plants acted as-if they had intentionality, despite not having a designer, even in the metaphorical sense supplied by evolution, and despite there being nobody around to interpret them.[[12]](#footnote-12) This is because some systems “impose strong natural limits on their possible semantical interpretation” (Dennett 1990, p.53). Dennett observes that photographs and maps can have quite obvious and natural intentional objects. But even ones lacking etiological grounding (such as the first mutant organisms) are interpretable given certain constraints about function and optimality, such as dispositional fittedness to an environment. Although a “swamp” (or Blotzmann) alarm clock has no content (intrinsic or otherwise) insofar as it is just a highly unlikely configuration of molecules, it is also interpretable under the assumption that its purpose is to tell time; this purchases access to powerful predictions and explanations that would otherwise be unavailable, even if it is a “zombie.”

Such interpretations are imposed by whoever is taking the stance, and the intentionality ascribed is merely derived. These would be cases of smuggling intentionality through the blockade of *etiological* isolation. However, if derived content can be pseudo-isolated, then perhaps this should shift our intuitions about the *structural* isolation Blinky’s conscious experience supposedly tolerates.

Inside the thought experiment, Blinky is alone, and yet she is still interpretable (and interpreted) by the person “outside.” Metaphysical isolation tests can be compromised like that generally. Another illustrative case is the employment of isolation tests in environmental ethics. Some think these can yield metaphysical intuitions about the intrinsic value of non-sentient life (Hill, 1983). But that I judge a world with only non-sentient life to be better than a world with no life at all is no proof of the intrinsic value of non-sentient life. It only shows something about my preferences and valuations. The putative intrinsic value could really just be explained by covert desire (covert in that the thought experiment only pretends to eliminate the evaluator and her preferences).

Similarly, could what looks like intrinsic intentionality come to be just more attitude-smuggling? A kind of derived intentionality is compatible with metaphysical isolation, and seem to be nonderived, if we miss the fact that an intentional framework is being deployed. In visualizing Blinky’s point of view one is projecting one’s own first person perspective into the thought experiment—I think of my own whiteness-qualia in considering what *e* is like for Blinky. But, then isn’t it still plausible that my own networked and derived self-interpretations are framing the attribution of consciousness and intentionality to Blinky? Though I admit to being unsure, this seems a very live possibility.

References

Anderson, H.K. and Grush, R. (2009). “A brief history of time consciousness: Historical precursors to James and Husserl,” *Journal of the History of Philosophy* 47(2): 277-307.

Bayne, T. (2010). *The Unity of Consciousness*. New York: Oxford University Press.

Bourget, D. (2010). “Consciousness is underived intentionality,” *Noûs* 44 (1): 32-58.

Burge, T. (1979). “Individualism and the mental,” *Midwest studies in philosophy* 4: 73-122.

Coren, S., Ward, L. and Enns, J. (2004). *Sensation and perception* (6th ed.). New York: Wiley & sons.

Dainton, B. (2010). “Temporal consciousness,” *Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/entries/consciousness-temporal/> *accessed on 11/8/13*

Dennett, D.C. (1988). “ Evolution, error and intentionality,” in Y. Wilks and D. Partridge (eds.), *Sourcebook on the foundations of artificial intelligence* (pp.190-211). New Mexico University Press.

Dennett, D.C. (1990). “The myth of original intentionality,” in K.A. Mohyeldin Said, W.H. Newton-Smith, R. Viale, K.V. Wilkes (eds.), *Modeling the mind* (pp.43-62). Oxford: Oxford University Press.

Gray, J. (2002). “The sound of one hand clapping,” *Psyche* 8 (11).

Hill, T.E. (1983). “Ideals of human excellence and preserving natural environments,” *Environmental ethics* 5: 211-24.

James, W. (1890). *The principles of psychology*. New York: Dover.

Putnam, H. (1975). “The meaning of ‘meaning’,” *Minnesota studies in the philosophy of science* 7: 131-93.

Robbins, P. (2009). “Modularity of mind,” *Stanford encyclopedia of philosophy*.

<http://plato.stanford.edu/entries/modularity-mind/> *accessed on 11/8/13*

Van Essen, D.C. and Deyoe, E.A. (1995). “Concurrent processing in the primate visual cortex,” in M.S. Gazzaniga (ed.), *The cognitive neurosciences*. Cambridge: The MIT Press.

1. The PURE theory is compatible with a variety of metaphysical outlooks, including physicalism (2010, p.54). [↑](#footnote-ref-1)
2. I am assuming that a succession of type-identical states is on a par with a token instantaneous state. [↑](#footnote-ref-2)
3. This is contrasted with the idea of a durationless “strict” present. [↑](#footnote-ref-3)
4. For more on these and other accounts of temporal consciousness see Dainton (2010). [↑](#footnote-ref-4)
5. In addition, there might be further types of content to given that t2 is not just another experience happening later, but is experienced as the *successor* of t1. [↑](#footnote-ref-5)
6. What about dispositional derivation? This would follow if undergoing the first in the series disposes one to undergo more. Conscious subjects will be so disposed. [↑](#footnote-ref-6)
7. Statement 5 is interpreted as what Bourget calls “Minimal Functionalism,” namely the view that it is at least “*partly* constitutive of attitudes” that they bear “certain relations to distinct intentional states,” especially “complex dispositions to token certain patterns of mental states (phenomenal states and other attitudes)” (p.42). [↑](#footnote-ref-7)
8. Thoughts may also dispose one to token conscious feelings, e.g. of conviction, or, the pull of desire, as Bourget notes (2010, p.42). [↑](#footnote-ref-8)
9. That is, the manner in which “encapsulated” mechanisms operate independently of each other, much as the radio of my car is encapsulated with respect to the headlights; in other words, the functioning of the module does not take any account of processing beyond its own boundaries [↑](#footnote-ref-9)
10. It could be argued that removing even a single coin does change the causal role of other quarters (since they would become slightly rarer), whereas the absence of any single experience makes no difference at all to the character of any other experiences. However, even in examples of perceptual dissociation, the remaining causal roles are slightly different. [↑](#footnote-ref-10)
11. But this is only to a degree, e.g. damage to the visual cortex in the area known as V1 can result in a catastrophic loss of visual consciousness, leaving a residue of action-oriented non-conscious vision. [↑](#footnote-ref-11)
12. Dennett (1988, fn.9) credits Sober for noticing a “first mutant” problem for naturalized teleology. [↑](#footnote-ref-12)