

# Swampman's revenge: squabbles among the representationalists

FRED ADAMS & LAURA DIETRICH

---

---

**ABSTRACT** *There are both externalist and internalist theories of the phenomenal content of conscious experiences. Externalists like Dretske and Tye treat the phenomenal content of conscious states as representations of external properties (and events). Internalists think that phenomenal conscious states are reducible to electrochemical states of the brain in the style of the type-type identity theory. In this paper, we side with the representationalists and visit a dispute between them over the test case of Swampman. Does Swampman have conscious phenomenal states or not? Dretske and Tye disagree on this issue. We try to settle the dispute in favor of Dretske's theory (to our own surprise).*

## 1. Introduction

There are at least two types of theories about the origin of the qualitative character of conscious experiences. There are the *internalists* who hold that the qualitative character of conscious experience is reducible to the electrochemical nature of activation in the sensory regions of the brain. For many philosophers who found the type-type identity theory of the mind to be plausible, perhaps the most plausible cases of identity were the identity of experiences with electrochemical activities in the brain. So for an internalist to explain why, say, pains feel the way they do one would point to the electrochemical nature of the pain centers of the brain. On this view, accounting for the qualitative feeling of the experience of pain would be exhausted by citing the electrochemical properties of the brain that correlate with being in pain. At least, this would be true if there is no genuine *explanatory gap* (Levine, 2001) in such an explanation. Whether there is such a gap is something we won't address here—but we don't think there is on the right sort of physicalism (Tye, 2001).

In addition, there are the *representationalists* who hold that the qualitative character of conscious experience is not reducible to the electrochemical activities (alone) in the brain. For the representationalists, the qualitative character of conscious experiences derives, at least in large part, from what is being represented (not from properties of the representing medium). So pains, for instance, derive their qualitative character from being representations of bodily damage (not from the

*Fred Adams & Laura Dietrich, Department of Philosophy, University of Delaware–Newark, DE 19716, USA, email: fa@udel.edu; laurann@udel.edu*

non-representational electrochemical properties of the brain). We would classify Block (forthcoming) and Searle (1983) as internalists, and Dretske (1995) and Tye (1995, 2002) as representationalists.

We should point out that there are also views that fall somewhere in between what we have called the views of internalists and representationalists. Views like that of Georges Rey (1997) treat qualia as tied to narrow content reports. We will be unable to go into detail comparing this type of view with those of Tye and Dretske in this paper. Here we will be mainly concerned with a specific dispute between externalist representationalist theories [1].

A test case for the debate between internalists and representationalists has been the case of Donald Davidson's (1987) thought experiment about Swampman. The idea of Swampman is to conceive of a possibility of an individual that is materially identical to a human (such as Davidson), but a being that has no prior history of involvement with an environment. Davidson asks us to imagine that lightning hits a stump in the swamp and fantastically a being qualitatively identical to Davidson materializes. (If this seems too fantastic, we suggest imagining a cloned human brain, conditioned in the lab to be physically identical to Davidson's brain but somehow with no sensory or cognitive involvement with its environment).

The reason this is a test case for the internalists vs. the representationalists is that, in principle, it should be possible for the swamp man's brain not to represent anything. If so, it should be possible for the swamp man to be a zombie—something with no conscious experiences. However, for an internalist, if the swamp man's brain is in a state qualitatively identical to Davidson's and Davidson is experiencing sugar in his mouth then the swamp man (with sugar in his mouth) should be experiencing the same thing as Davidson.

As we say, in principle, it is possible for the representationalist to deny that the swamp man experiences anything at all. Indeed, at least one representationalist (Dretske, 1995) says just this. Another representationalist (Tye, 1995, 2002) says that a swamp man would have conscious experiences, if he meets the general conditions for Tye's representationalist account of qualitative content.

We intend to declare a winner in this broader debate. We find ourselves leaning towards representationalism. Still, those so leaning have two views from which to choose. What we will do here is detail the choices. That is, we will explain the differences among the representationalists and how they arrive at those differences. Then we will evaluate the strengths and weaknesses of the two positions within the representationalists' camp. At the end we surprise (even ourselves) in declaring a winner—for now.

## 2. Dretske's view

We will begin with the view of Fred Dretske (1995). Dretske defends the hard line that Swamp Fred would have no conscious, phenomenal states. Dretske believes this because his view of consciousness requires that regions of the brain acquire the cognitive function of being indicators. They indicate instances of properties in the environment (or body). A consciousness of sweetness in Fred derives from the part

of Fred's brain that has the job of indicating sugar (Fred has sugar in his mouth). A part of Fred's brain acquires the job of indicating sugar by being selected for its ability to do so. No parts of Swamp Fred's brain have cognitive functions (according to Dretske). We say more about this below.

On Dretske's view, sugar's causing activation in the brain alone is not sufficient for a conscious experience of sweetness. For conscious experience to arise, the parts of the brain that give rise to the experience must not only fire when and because of sugar, not only indicate the presence of sugar, but have acquired the function of indicating sugar. This happens when the relevant parts of the brain that indicate the presence of sugar have been recruited or selected for the ability to track the presence of sugar.

Although Dretske uses the term "function" in his theory, we don't believe too much weight should be placed upon that term (Adams, 2003). What is essential is that there is a type of *sustained causing* (Adams, 1991). That is, the structure in the brain  $s$  that indicates the presence of sugar will cause some other brain activity or bodily movement  $m$  (say, *swallowing*). When the  $s$  causes  $m$  (rather than some contrasting  $n$ —*spitting*) because of the indication of sugar by  $s$ , then and only then does  $s$  acquire the function of indicating sugar. What is important is the sustained, contrastive causing. The structure  $s$  must be sustained in its causing some relevant effect by the fact that it indicates the presence of sugar. If so, then  $s$  comes to represent the presence of sugar and makes Fred (in whom  $s$  does this) conscious of sugar by virtue of the qualitative experience of sweetness (in the mouth).

We will say more below about whether acquisition of function by a structure like  $s$  can take place within an individual's lifetime or whether such functions are inherited due to selection across individuals. For now, we will just say that on Dretske's theory, when Swamp Fred instantaneously arrives on the scene, he has no history and has not been around long enough for cognitive functions to have formed. He materialized with no selectional history or history of intracranial recruitment. Therefore, even if sugar in Swamp Fred's mouth causes activation of region  $s'$  of his brain that is neurophysically identical with region  $s$  in Fred's brain,  $s'$  has no cognitive function. Nothing in Swamp Fred has the job of alerting him to the presence of sugar (in the mouth), and Swamp Fred experiences no sweetness (or anything else, for that matter). As Dretske says, "Everything is 'dark' in the mind of Swamp Fred" (1995, p. 145).

### 3. Tye's view

Tye (1995, 2002), like Dretske, is a representationalist. For him the phenomenal content of consciousness derives from the representational properties of the brain. In principle, we take his view to acknowledge the possibility that all *might be* dark inside the head of Swamp Tye. This would happen only if no brain structures in Swamp Tye represent properties in his body or environment [2]. In principle this should be possible on any representationalist account of phenomenal content: no representation—no content. Not every structure physically identical with another structure that is a representation is equally a representation.

We offer confirmation for this claim in Tye's own words. At one point he says:

The thesis that it is metaphysically possible that there are microphysical twins that differ with respect to the phenomenal character of their inner states (a thesis to which the representationalist is committed if there can be wide phenomenal representation) comports well with the thesis that it is conceptually possible that there are microphysical duplicates that differ phenomenally. (2002, p. 65)

We take this to mean not only that there may be microphysical twins that have different phenomenal content, but that one of the twins may *differ* in having *no phenomenal content*. At another point Tye says, "Intuitively, cases of absent qualia are also epistemically possible. It could conceivably turn out that, in some actual people, *no* phenomenal state occupies the role S occupies in you." If this can happen for one state S in you, why not for all in your microphysical twin?

All of this is consistent with Tye's other claims: "... phenomenology *ain't in the head*," "To discover what it is like, you need to look outside the head to what the brain states represent," "... systems that are internally physically identical do not *have* to be phenomenally identical" (1995, p. 151). In general, then, it seems that Tye is committed to the in principle possibility of all being dark inside Swamp Tye. Whether this is actually so or not would depend on whether there were representations inside.

Despite this in principle possibility, Tye steadfastly maintains that *there will be* representations inside the head of Swamp Tye, when he actually comes into existence. He maintains this because he defends two different versions of representationalism, both of which permit Swamp Tye to have phenomenal consciousness. We will explain why Tye changes his view of representation from a straight co-variation view to an asymmetrical dependency view (of the type defended by Fodor, 1990), but on either view Swamp Tye would have experiences—contrary to Dretske's view.

Tye (1995) begins defending a representational theory of the phenomenal mind on the basis of a co-variation view of representation. The view states that a structure S represents that P = df. If optimal conditions obtain, S is tokened in x if and only if P and because P. In this definition, "X" is a place-holder for a person and "P" is a place-holder for a proposition (due to the "that" on the left hand side of the identity sign). We think the latter is a particularly bad choice because of two factors. First, propositions don't cause things. Events or instantiations of states of affairs may cause things and may have propositional content, and may cause things because they have propositional content, but propositions themselves don't cause things. Second, Tye himself says the structure of phenomenal representation is topographical (1995, p. 120). To us this suggests that it would be far better to interpret "P" in the definition as a property instance. If so, then the definition says that S (sensation of sweetness) represents P (sugar in the mouth) if, under optimal conditions, S is tokened in person x if and only if there is sugar in the mouth and because there is sugar in the mouth. Hence, the sweet sensation represents sugar (in the mouth) because it is *tokened when there is* and *caused by* sugar (in the mouth).

Clearly, the contrast between Tye and Dretske on representation is at the heart of the difference in their views of a swamp man. There is nothing in Tye's co-variation account of representation that requires a causal history or acquisition of cognitive function prior to the event of a swamp man coming to have sugar in his mouth, and thereby coming to have a sweet sensation. What we shall do in the remainder of the paper is evaluate the support for the competing views of representation being offered. The different theories of representation yield the differences of position on whether a swamp man would have phenomenal consciousness (initially).

Before we can proceed with this evaluation, we need to point out that Tye's view has changed. At the end of a long set of objections and replies to his theory, Tye (2002) points out a problem of his own that he admits requires a change of view. He says:

Suppose that sensory state *S* causally covaries with perceptible quality *P* under optimal conditions. Suppose, moreover, that *P* is nomically correlated with imperceptible quality *Q*. Then, according to the presented account, sensory state *S* represents not just *P* but also *Q*. However, if *Q* is imperceptible, it cannot contribute directly to the phenomenal look of a thing. And that spells trouble for the representationalist. (2002, p. 139)

First, notice that Tye has switched from talking about *P* and *Q* as propositions to talking about them as property instances, as we recommended. Second, Tye's solution to this problem is to adopt a Fodorian asymmetrical dependency view on which, as long as it is true that *Q* causes (and correlates with) *S* *because* *P* causes and correlates with *S*, but not vice versa, then *S* represents *P*, not *Q*. The connection between *Q* and *S* piggybacks on the connection between *P* and *S*, on such a view, and that allows *S* to univocally represent *P* (not *Q*, nor *P* or *Q*, nor *P* & *Q*) (Adams, 2003; Fodor, 1990). Notice that Tye drops the earlier "optimality" clause of his prior co-variation account. There is no "optimality" clause in Fodor's version of asymmetrical dependency, and whether this difference is significant will become clear below as we evaluate Tye's and Dretske's views.

#### 4. Objections to Dretske's view

Objections to Dretske's view come in two types: epistemological and metaphysical. Under the epistemological, there is the worry that for all we know we might be swamp people, on Dretske's theory. But that is preposterous. So, the theory must be false. Under the metaphysical, there are two worries: (1) why does causal history matter? and (2) won't the mental be epiphenomenal, if mental differences don't show up in physical (and, hence causal) differences? We take these in turn.

For the epistemological objection, we will follow Tye's version (2002, Chapter 6), although similar points will be found in Levine (2001, pp. 113–119). Tye recounts a reply, on Dretske's behalf, to Block's "Inverted Earth" (Block, forthcoming) example, and notes that such a reply has an important negative epistemological consequence. Block gives the example of someone (say, Tye) going from Earth to Inverted Earth. On Earth things have their normal colors and looks. On

Inverted Earth, things have inverted spectral colors (and presumably inverted looks). So on Inverted Earth, the sky is yellow and grass is red, and so on. Block objects to representationalist views of phenomenal content by claiming that if Tye were transported to Inverted Earth with inverting lenses, then the sky would look blue to Tye, but his brain would be representing a yellow sky. Hence, the representationalist cannot account for the blue look by appealing to the color the brain actually represents (yellow).

On Dretske's behalf (and correctly we maintain), Tye offers the following type of reply. Tye's experience is *of blue* not because his experience is now being caused *by blue* or is now *representing blue* (or *yellow, for that matter*), but because it is a type of experience that has acquired the *cognitive function of indicating blue*. Though this brain state is now being triggered by a yellow sky (and inverting lenses), Tye's experience is of blue because of his (or his ancestor's) past history with blue (during which cognitive function formed). This reply nicely handles Block's case and is perfectly consistent with the representationalist theory of phenomenal content. It explains why something not blue might look blue. It also explains why physically identical individuals (Earth Tye and Inverted Earth Tye) might be in the same physical states and be experiencing different qualia. It is because Earth Tye is looking at a blue sky (no inverting lenses) and Inverted Earth Tye is looking at a yellow sky (no inverting lenses). Their similar brain states have acquired different sensory indicator functions.

Tye's first objection to this response is that it is "highly counterintuitive," when applied to a swamp man. For Swamp Tye's brain states, unlike Inverted Earth Tye's, would have *no sensory indicator functions* (not just different indicator functions). Swamp Tye has no history with an environment and no ancestors with a history (to pass on inherited indicator functions). Hence, what Tye finds counterintuitive is that Swamp Tye would have no phenomenal mind at all, even though he was Tye's physical duplicate.

We don't place too much weight on the claimed finding of "counterintuitiveness." On Tye's view too, as we claimed above, if there were no *representations* in Swamp Tye, he would lack a phenomenal mind. Physical identity does not guarantee representational identity, even for Tye. What would guarantee representational identity would be satisfaction of the conditions on sensory representation—and what those conditions are is what is now in dispute. We return to this below.

Tye's second objection is that if Swamp Tye really lacked phenomenal consciousness, then, in principle, it would be possible for Tye or Dretske or anyone to discover that they were swamp people. Tye urges, "... patently that is as ridiculous for Dretske as it would be for you or me" (Tye, 2002, p. 119) [3].

This second objection goes by pretty quickly. Its correctness is supposed to be readily apparent, but it is less apparent if unpacked more slowly. To begin, for Swamp Tye (or anyone) to discover he was a swamp person, he would have to be able to think. On a view like Dretske's, there is no reason to maintain that Swamp Tye could have beliefs or thoughts. After all, Dretske is an indicator functionalist about thoughts and beliefs too (Dretske, 1988). The instant Swamp Tye material-

izes nothing in his brain (according to Dretske) has meaning (conceptual indicator function). So the instant Swamp Tye materializes, he would not be able to discover that he was a swamp person.

Now, interestingly, to get the “ridiculousness” Tye wants in his argument, he suggests that, in time, Swamp Tye should be able to acquire conceptual indicator functions (i.e. beliefs). And if he cannot also acquire phenomenal property indicator functions (experiences), then he should be able to discover that he is a swamp person. He should be able to discover that he has beliefs but no experiences. This is what Tye finds to be *ridiculous*.

Again, we will move slowly. To begin, Tye himself thinks that concepts of phenomenal states are what he terms “phenomenal concepts.” In Frank Jackson’s Mary case, Tye explains that “She does not know what it is like to experience red; and intuitively knowing what it is like to have that experience is necessary for possession of the phenomenal concept *red*” (2002, p. 17). Mary could not know what she was missing, not having experienced red. For to know what she was missing, she would have to experience red phenomenally and then acquire the concept of that phenomenal content. Barring that, *Mary really didn’t know what she was missing*, prior to experiencing red and acquiring the phenomenal concept of that experience. Similarly, we submit that Swamp Tye literally would not know what he was missing. So he would not be able to discover that he was missing something, contrary to Tye’s suggestion to the contrary. For Swamp Tye would possess no phenomenal concepts (if Dretske is right that he lacks all phenomenal consciousness).

So where is the “ridiculousness?” What is ridiculous is that someone who has phenomenal concepts (Tye, Dretske, or anyone reading this) might discover that they are swamp people. That is ridiculous because you (they) already have the relevant phenomenal concepts, and the only way you (they) could acquire them is by having the relevant phenomenal experiences. What is also ridiculous is that someone without phenomenal concepts might discover that they lack phenomenal contents, but we submit that this is not possible (epistemically) for swamp people to do. So it cannot be an objection to Dretske’s theory (for his theory does not permit it either).

Furthermore, if we suppose that a swamp man can acquire concepts (conceptual indicator functions), we may also suppose that a swamp man would be able to acquire experiences (non-conceptual indicator functions). It is true that Dretske (1995) suggested that sensory functions might be acquired in one’s developmental inheritance but conceptual functions may be acquired in an individual’s lifetime, and that difference is what inspired Tye’s suggestion that Swampman might engage in self-discovery about himself, on Dretske’s theory. However, Dretske no longer feels the need to say that one cannot acquire sensory functions in one’s lifetime (in personal correspondence). However, he still thinks one would need to explain how those functions are acquired and how sensory indicator functions differ from conceptual indicator functions (Dretske, 1981). But in both cases, such acquisition would take time and not happen the instant the swamp man formed. If at the instant the swamp man formed, he lacked phenomenal concepts and phenomenal experi-

ences, because he had no indicator functions of any kind, there is nothing *ridiculous* in that. It may be false, but not ridiculous. Whether it is false is what is in dispute between Dretske and Tye.

We can imagine that Swamp Tye learns that he is a swamp man from someone else who has phenomenal concepts (Dretske, say, who just saw Swamp Tye come into existence). We are still not sure what Swamp Tye would learn—lacking the phenomenal concepts—but if he could learn that he was a swamp person in this way, we don't see anything ridiculous in the idea of this self-discovery.

Lastly, and interestingly, Dretske accepts that one might know what one thinks, but not that one thinks it (Dretske, forthcoming a). We suspect that he would not find it ridiculous that one might have introspective access to a physical structure in the brain that is an experience but not know it is an experience (not know that it is an experience by the same introspective route of access to the experience). Indeed, Dretske (forthcoming b) maintains that we do know that we are not zombies, but he maintains that it is harder than Tye thinks to explain *how* any of us know that we are not zombies (and hence, not swamp people).

Here is our take on some of this. Suppose we could be shrunk and inserted into Tye's brain, as in the Disney film *Incredible Voyage*. In a miniature vessel like a submarine, we might come upon a structure in the brain that is an experience of blue (that is firing in response to Tye's looking at the blue sky and that has the function of indicating blue, à la Dretske's theory). Our access to the structure is physical. We are right there in the brain observing it via our special imaging device. We are observing something that is an experience, but we are not observing its being an experience. To observe that, we would need to observe its being a structure with an indicator function. But having an indicator function is not something that can be read off its physical properties (two things can be physically type identical but functionally non-identical). Nor can something's having a function be observed in a single time-slice. Acquiring a function takes time, more time than the duration of our observation of the activation of neural structure in Tye's brain. So, in a sense, we would observe a structure that was an experience of blue, but not observe its being an experience of blue.

When Tye introspects, some of his own brain's internal detectors go proxy for us in our miniature submarine. Tye's own neural detection mechanisms would have physical access to the structures in his brain that constitute his experience of blue and that represented blue and had the function of indicating blue. But his brain would also be taking a snapshot in time of this neural activation. His detection mechanisms would be no better than ours at telling him—of this experience of blue—that *it is an experience of blue*. His introspective detection mechanisms too give him only a time-slice of the activities of his neural structures. They do not reveal whether those structures have indicator functions. (For that matter, the internal scanners do not reveal whether Tye's own conditions of representation are met ... correlation between properties under ideal conditions or asymmetrical causal dependency.)

Here is what Dretske says about the prospects of one's being able to tell, by introspection, that one is conscious (consciously experiencing a blue sky):

... If a conscious experience in S is some electrical-chemical state of S's brain, then if S is aware (introspectively or otherwise) of that experience, S is aware of that electrical-chemical state of his brain. And *vice versa*. A neurosurgeon observing this electrical-chemical state of S's brain *sees* the object that S "sees" by introspection—S's conscious experience. Can this be right? Does anyone think we come to know we are conscious by inspecting our own brains? Even if we could inspect our own brain, what is it about our brain that would tell us we were conscious? Would it also tell the neurosurgeon who sees our conscious experience that we are conscious? How? What electrical-chemical property of a brain state testifies to a brain state's being conscious? ... If we know we are not zombies, that doesn't appear to be the way we know it. (Dretske, forthcoming b, p. 13)

Once again, our take on this is that Dretske thinks introspection (i.e. brain scanning) won't tell someone (anyone) whether the state introspected has an indicator function. And if being an experience is its having an indicator function, simply scanning the physical properties of a brain state S (which is a conscious experience) won't reveal the external components of being a conscious experience (that it has an indicator function).

On this way of thinking, Tye's and Swamp Tye's brains would be running physically identical scanners. And if their scanners alone would not reveal that the scanned state S (in Tye) or S' (in Swamp Tye) were (or were not) experiences, then it would be harder than Tye thinks for Swamp Tye (or Tye) to discover that he was (was not) a swamp person.

We have already responded to Tye's claim that, on Dretske's theory, one might discover one was a swamp person. Clearly, Dretske does not think that the problem of knowing one is not a swamp person (zombie) is ridiculous. He thinks explaining how one knows one is not a zombie is a significant challenge—yet to be met. With respect to Dretske's challenge, we believe that the scanning mechanisms in conscious beings and zombies must work differently. In conscious beings, when the introspective scanner hits the phenomenally conscious state S (a state of experiencing sweetness), the diaphanous nature of conscious experience will reveal to introspection the sweetness being experienced. That is, the state of the brain that has the function of indicating the presence of sweetness (sugar) will become activated. Since this state has the function of informing the brain of the presence of sweetness (sugar), and since it acquires its qualitative phenomenal character from the sweet sugar itself, its activation brings with it the phenomenal sweet that we associate with sugar. The phenomenal character consists of the objective properties of the sugar being made present to the conscious mind via activation of the indicator function of the brain. If Tye has sugar in his mouth, experiences it phenomenally, and introspects, his introspective scanner will make him aware of his being aware of sweetness (in his mouth). His introspection may reveal both a re-affirmation of the (non-conceptual) *experience* of sweetness and a *belief* that he is experiencing sweetness (in which he applies the phenomenal concept of sweetness to his experience). Whereas, we suspect that in Swamp Tye (if a zombie), his internal scanner would do neither

of these things when it arrives at the internal state *S'*. *S'* in Swamp Tye is the physical duplicate of the qualitatively conscious state in Tye, but in Swamp Tye is a state absent of qualia. The state in Swamp Tye will be absent of qualia because, although it is caused by the sugar in his mouth, it is not a state that has acquired the function of indicating the sweetness of the external substance to the brain. There will, therefore, be no representational character of the sugar made present to the brain of Swamp Tye. We have already indicated why, lacking phenomenal concepts, Swamp Tye would literally not know what he was missing when his internal scanner arrives at state *S'*. This is why he would not discover that he was a zombie by internal scanning alone.

Now let us look at the metaphysical objections to Dretske's form of representationalism. We will take both objections together. (1) Why does causal history matter? (2) Won't the mental be epiphenomenal, if mental differences don't show up in physical (and, hence causal) differences? The basis for the metaphysical objection is that history is not a causal power. It should not matter to the causal powers of a brain how the brain got the way it is. What it can cause is a function of its physical properties. So if there could be two physically identical brains, one conscious, one not, how could being conscious matter? How could the conscious brain cause something the non-conscious one couldn't (if both were physically identical and, thereby, had the same causal powers)?

We suggest that the case of a swamp man clouds the issue of why causal history matters and why consciousness is not epiphenomenal. It clouds the issue because the real contrast is between a brain that wasn't miraculously formed to be qualitatively identical with Davidson's and Davidson's—not between Davidson's brain and Swamp Davidson's. By that we mean that the swamp man case sets up an artificial contrast. (We see Fodor, 1994, p. 38, make a similar point about Twin-Earth examples, namely, that their nomological implausibility allows them to be safely ignored when looking at causal powers.) In nature, there are no swamp people. And it is in nature that causal history and consciousness matter.

To see this, let's compare another contrast. What is the value of knowledge over justified true belief? If truth is the value of knowledge, then having justified true belief gets you truth every time knowledge does. So how could one (knowledge) be more valuable than the other (justified true belief)? An answer (Dretske, 1989) to this question that we find plausible is that knowledge is superior because the mechanisms that underwrite knowledge are superior (reliable cognitive processes). The mechanisms that deliver truth in cases of justified true belief are capable of (and do deliver) false beliefs on other occasions. The mechanisms that deliver knowledge, in contrast, deliver false beliefs not at all (or at least not as often—depending upon one's view about knowledge). The point is that the relevant contrast is not truth that comes from knowledge producing mechanisms vs. truth that comes from justification producing mechanisms. The relevant contrast is between knowledge producing vs. non-knowledge producing mechanisms themselves. The fact that both can on occasion deliver truth clouds the difference in their cognitive values.

We think the same thing is going on in the case of a swamp man. The brains of Davidson and Swamp Davidson are artificially contrived to be identical (in the

way that justified belief and knowledge can both deliver truth). They are contrived by the stipulation of the example to be identical and have the same causal powers. However, in the wild (as it were) brains don't get wired like Davidson's unless one is conscious of one's environment. And one is not conscious of one's environment unless there are internal structures that correlate correctly with what is happening in one's environment. On Dretske's view, to be sustained, that correlation requires indicator functions. So a brain that is conscious is wired differently from one that is not (outside of contrived stipulations). And a brain that is conscious has different causal powers than one that is not because of its wiring, as it were. Its wiring is different because of its causal history. So on this view, the proper contrast is not Davidson's brain vs. Swamp Davidson's, when one is asking why history matters or whether consciousness is epiphenomenal. The proper contrast is Davidson's brain awake vs. Davidson's brain sleepwalking (or some similar non-contrived contrast). Here we would see why consciousness and causal history matter (they keep Davidson from bumping into walls or falling down stairs). One coordinates activities with contingencies of one's environment by representing and updating internal representations of the environment. A conscious brain can do this and the proper representations give one consciousness of one's self and surrounds.

The case of a swamp man is a contrivance. If the swamp man swallows when sugar is in his mouth, the swallowing need not occur because it represents (experiences) a sweet substance. It is true that there will be a neural structure there that is physically identical with that in Davidson (that which in Davidson is an experience of sweetness). But that there is parallelism is the contrivance. It might be a metaphysical possibility, but there is no reason to think that such a brain *would* form even if it *could*. That Swamp Davidson's bodily movements correlate with events in his environment is *sheer accident ... sheer stipulation*. However, in a conscious brain the coordination is *not accidental at all*. That is what consciousness is for—solving this coordination problem.

So why does causal history matter? Because without the right history, the brain would not wire itself in such a way as to be able to represent and update representations of its environment. The brain must causally interact with the relevant properties in one's environment in order to acquire indicator functions. The relevant wiring of the brain constitutes the acquisition of these indicator functions. And why isn't consciousness epiphenomenal? Because being conscious of something like sugar in the mouth *just is* having the property of sugar's presence presented to the mind (brain) via the structure whose function it is to so indicate. Being conscious of the sugar is the indication's taking place in the mind. When one is phenomenally conscious of the sugar, then one may take appropriate action (swallowing, seeking more sugar or food) *because one is conscious of the sugar*. Without consciousness (of sugar, food), the brain would not be wired in the way that it is for the reason that it is not producing and would not produce behavior that non-accidentally coordinates with the organisms needs and the changes in its environment. Minds are for solving this coordination problem. If so, why doesn't a swamp man need a mind to solve the coordination problem? Because the contrived example solves it for him—

by stipulation. If all the world's people were swamp people, there would no need for conscious minds. Of course, none of the world's people are or will be swamp people.

Lastly, we would point out that there is a difference between what Dretske (1988) elsewhere calls "triggering" causes and "structuring" causes. Say that a state of Swamp Tye's brain (A) causes (B) his swallowing. We might say that A is a triggering cause of B. But Dretske would ask the further question, why does A cause B? To answer this is to give what he calls a "structuring" cause. If A causes B because A has the function to represent sugar (and is therefore phenomenally sweet in Tye, but not in Swamp Tye), then there is a difference in structuring cause in Tye that is not present in Swamp Tye. This is true even though there is no difference in triggering cause (they both have brain state A). So we believe that there is a relevant metaphysical difference in cause, but one must look to the explanation of why states like A cause states like B to detect it (Adams *et al.*, 1990).

We have discussed the two sorts of objections to Dretske's account. Now we turn to objections to Tye's account.

### 5. Objections to Tye's view

As we pointed out in Section 3, Tye has two versions of his representationalist theory: the correlation version and the asymmetrical dependence version. We take the latter to be his considered view, since he adopts it in response to his own objections. But we will begin with his correlation version.

On the correlation version, Tye (1995, p. 154) endorses a swamp man's having phenomenal consciousness.

... on the causal covariation model of representation, no obvious difficulty arises for the claim that some of Swampman's inner states represent things. States in his head certainly track various external environmental states, just as mine do. Moreover, given that there are no distorting mirrors, no special peculiarities in his environment, his behavior is entirely appropriate to the states that are tracked ... So it is natural to suppose that, for a being of his sort, without any evolutionary history to weigh, optimal conditions obtain and hence that there is sensory representation of those external states.

First, we don't see how Tye's "optimal conditions" clause applies to swamp people. His account requires this clause, since he wants to explain such things as phantom limb phenomena (1995, p. 112). In such cases, it may feel as though one's left foot is being crushed, when one's left leg has been amputated at the knee. Since the left foot is gone, there is no property of its being crushed that correlates with and is causing the phenomenal mental representation. So Tye needs to invoke the clause that what is being represented is what would be the case under optimal conditions for such a person. But we believe optimal conditions are type identified relative to types of individuals. It begs the question to assume that a swamp man is of the same type as we are (or Tye is). In fact, it is not even clear that he would be human. He would not be if there is an historical component to genetic identity (as there seems to be). Thus, it seems that Tye's "optimal conditions" for his swamp man are

*undefined*. As such, we fail to see how Tye can appeal to them in explaining what internal states of his swamp man represent. It is true that states in the swamp man's environment would cause chemical-electrical states in his brain. But not everything B that is caused by something A is a representation of the cause A. So the mere facts cited in the quotation above about inner correlations with outside conditions (and resultant bodily motions) would not guarantee that genuine phenomenal representations would be taking place inside the swamp man.

Second, we believe Tye would have his own Twin-Earth problem. He admits that his physical twin would have phenomenally different states if correlated with different properties in a Twin-Earth type setting. "It is still possible that two different organisms that evolved in different ways, while nonetheless sharing the same internal microphysical states (at some given time *t*), differ in their phenomenal states at *t*" (Tye, 1995, p. 153). So we would press the question, if Twin-Tye's internal representation is representing property P on Twin-Earth and Tye's microphysically identical representation is representing property Q on Earth, what property is Swamp Tye representing (P, Q, P or Q, neither, both)? And are the "optimal conditions" for Swamp Tye those of Tye, Twin-Tye, neither, both? It seems to us that there is no clear answer in Tye's correlation version of a theory of representation.

Now Tye may well be onto these sorts of concerns. Instead of appealing to "optimal conditions," in his recent book he appeals instead to *conditions of well functioning*.

This leads to the thought that Swampman can have inner states that acquire representational content via the tracking or causal covariation that takes place under conditions of *well functioning*. However this is spelled out—whether or not, for example, it is demanded that he become a full-fledged member of some appropriate community—the general suggestion is that, where the representational contents of experiences are concerned, what counts in normal conditions can vary with the kind of creature ... Where there is a design, normal conditions are those in which the creature or system was designed to operate. Where there is no design, normal conditions are, more broadly, those in which the creature or system happens to be located or settled, if it is functioning well ... in that environment. (Tye, 2001, p. 122)

This quote is clearly intended to answer the sorts of objection we raised to his earlier account. He is replacing "optimal conditions" with "normal conditions" and "conditions of well functioning." And he seems to be doing this precisely because it is not clear what would count as "optimal" for a creature "not designed" by nature or anyone.

Unfortunately, in this we see change, but not progress. For all the old problems about relativizing optimal conditions to types of individuals still apply when relativizing normal conditions or conditions of well functioning. Relative to whom is Swampman operating under normal conditions or conditions of well functioning? If the answer is "relative to the type *swamp men*," then clearly Tye is begging the

question. And if that is not the answer, then what is? His theory still faces all of the open-ended choices it faced before when he appealed to “optimal conditions.” Again, we see no clear answer to the problems for Tye’s correlational account of representation, and we find it still unsatisfactory for these reasons.

We might add that the whole notion of a *function* seems to require something like a temporal process of selection for a beneficial effect or at least a sustained causing (Adams, 1979; Enc & Adams, 1998). It, of course, takes time for functions to form ... time Swampman doesn’t have, when first he materializes. We admit that functions may develop quickly within the lifetime of an individual, but still function acquisition takes time and nothing would have acquired functions in Swampman at the moment of materialization.

Furthermore, Tye seems to be aware that his account has the problem we note above about the relevance class for Swampman. Furthermore, he is aware that this presents something of a “mixed” view (a pinch of Tye and a twist of Dretske, blended together). For in attempts to avoid counterexamples to his view which pit puzzles about Swampman against Block’s puzzles about Inverted Earth, Tye says the following about his own appeals to normal conditions and well functioning:

This may initially seem to provide the representationalist with a route between the Swampman problem and the Inverted Earth example. Swamp duplicates may now be credited with experiences ...; moreover if you or I travel to Inverted Earth, the representational contents of our experiences remain the same in corresponding situations. Unfortunately, the path leads to a dead end, for if we can travel from Earth to Inverted Earth, so too, can swamp creatures. The case of the traveling swampman, equipped with inverting lenses, lies beyond the resources of the above mixed, representational theory. Here, representational content will change, but phenomenal character will remain the same. Representationalism, it seems, is in deep trouble. (Tye, 2001, p. 122)

Now where is the trouble? Tye is not specific about the exact nature of the “deep trouble.” If we are right so far, the trouble lies in identifying the relevance class into which to place Swampman in the Twin-Earth and Inverted Earth types of puzzle cases. Into which group or class should Swampman be placed? Which is his natural class? Tye seems to give in to the challenges and develop a different account—the asymmetrical dependency account. So it is to that account that we now turn. However, we can’t resist pointing out that even if one had a principle by which to decide upon a relevance class for Swampman, this would still ignore the more basic issue of why anything inside Swampman would have conditions of well functioning (given his non-history).

Since Tye now accepts an asymmetrical dependency account of representation, his account must face all of the problems such accounts face in general (Adams, 2003). Instead of recounting all of those problems here, we will cut to the chase and present the ones we think are most relevant, given what we have said above. On asymmetrical dependency accounts, a represented property instantiation of P causes representation R, and if any other property instantiation Q causes R, it does so only

because P causes R (but not vice versa). Notice the following. First, this type theory will face Twin-Earth puzzles. Suppose there is a type of substance on Twin-Earth (possibly aspartame) that tastes like sugar and is Twin-sugar. Now when Tye tastes sugar, it tastes sweet. When Twin-Tye tastes Twin-sugar, it tastes sweet. The asymmetrical theory gets wrong the explanation of the phenomenal sweetness. It must either say that Twin-sugar tastes sweet only because sugar does (which gets it wrong for Twin-Tye) or vice versa (which gets it wrong for Tye). And if it says it is a dependency on sugar or Twin-sugar causing the sweet taste, then it gets it wrong for both. For then the sweet tastes in the mouths of both represent the disjunction of sugar or Twin-sugar. We see no reason to think that when sugar tastes sweet to Tye that this has anything at all to do with Twin-sugar (especially if we were to build the case so that Earth were devoid of Twin-sugar). And if Tye claims that nothing tastes sweet to Tye unless sugar does, this is falsified by the fact that Twin-sugar tastes sweet to Twin-Tye (Tye's physical duplicate). So the claim of asymmetrical dependence seems plainly false. Hence, the theory seems to fail to be able to handle such Twin-cases.

Second, we believe the claims of asymmetrical dependency themselves are suspect. Why would there be any such dependencies at all? On teleological theories, tied to function acquisition, we can understand the origin of such dependencies. The function of the heart is to circulate the blood, not to make heart sounds. The heart makes heart sounds but only because it has the function to circulate the blood. There is clearly an asymmetry here, but it is explainable by the fact that hearts were selected for their function to circulate blood, not to make heart sounds. However, on the asymmetrical dependency theories of representations, asymmetries are simply stipulated as metaphysical truths. They are not explainable by facts such as selection (since Tye clearly rejects such teleological facts in embracing a phenomenology for Swampman). Thus, such theories have *no answer* to the question why there would be such asymmetries in the first place. Why would it be true that substance Q wouldn't cause sweet taste R unless sugar did? For example, why couldn't an electronic probe in the brain cause the very same stimulation that the sugar causes? The theory has to either deny that the probe could cause it or say that the probe could cause it only because the sugar can. But why would this be true? The theory has no answer but to stipulate these sorts of asymmetries as metaphysical bedrock. We find such claims to be highly implausible. There surely are basic laws—laws not further explained by other laws of a more basic science. But laws about what sugar can cause in the brain do not seem like the sorts of laws that can be basic. So from where do the asymmetrical laws about phenomenal facts come, if not from selective histories? Asymmetrical dependency theories have no answer.

Lastly, Tye accepts that there can be phantom limb pain in children who were born without limbs, and cites the relevant research to support this (Tye, 1995, p. 151). So he has something of an uninstantiated property problem. A representation R represents a property instantiation P, but in this case there is no property of having a limb to be instantiated. So how does R become locked to property P in order to represent P (having a limb)? On teleological theories, developmental history acquired from one's ancestors may explain phantom limb phenomena in children

without limbs. But Swampman has no ancestors. So a swamp duplicate of a child born without limbs experiencing phantom limb pain will, according to Tye, experience phantom limb pain. How? How is that possible? We don't see how the asymmetrical dependency theorist can give a plausible explanation. One would just have to stipulate that something Q would cause the representation R (of the limb) only if P (the limb) would (did) *even when the child has no limb!* Nice trick!

We are sure there are many other such problems that are direct analogs of the problems faced by other asymmetrical dependency theorists of meaning, but we will stop the objections here for now.

## 6. Conclusion

In this paper we have surveyed the differences between the two most prominent representational theories of phenomenal content. We have highlighted their differences by paying attention to what they would say about whether a swamp person would have phenomenal consciousness. We have considered objections and replies to both types of theory.

The only competing materialist theories are the internalists' theories which claim that it is the electrochemical properties of the brain alone that account for the phenomenal content of consciousness. We find these theories less plausible because of the problem of matching qualia with electrochemistry. There is vast array of qualia we distinguish. How do these qualia get matched to electrochemistry? The differences cannot be due to location in the brain alone. Different regions in the brain can assume different functions. It can hardly be due to differences in chemistry alone. There aren't that many different neurochemicals. It might be possible that it could be due to different electrical properties (spiking frequencies), but this flies in the face of externalists' intuitions about the mind that have caught hold over the last quarter of a century (intuitions we share and will not examine further here).

So that leaves us with a choice among the representationalists. At this point we find Dretske's theory more plausible than Tye's for the reasons we have given. For both of us, this represents a change of view from internalism about phenomenal content to representationalism. We used to believe that phenomenal content of conscious states had to be type-reducible to the electrochemical properties of the brain. So it is something of a surprise to us that we find ourselves coming over to the representationalists' way of thinking about phenomenal consciousness. And that is why we have focused on the squabbles among the representationalists.

We also surprise ourselves in lining up against the very powerful internalists' intuitions about the swamp man case. It seems incredible that someone could be in the exact same type of electrochemical state as someone else and not have the exact same phenomenal experiences. However, as representationalists we have to appeal to other cases where physical identity does not determine the relevant other kinds of identity. There can be counterfeit bills that are physical duplicates but not monetary duplicates. There can be photographic plates that are physically identical, but not representationally identical (because one was not exposed to light coming from a photographed subject, when the film was developed). Indeed, if one thinks, as we

do, that representing has to do with acquired function, then cases where two objects are physically but not functionally identical are the cases needed to blunt the internalists' intuitions. We are certain that there are many such cases in nature (human appendix now vs. in our distant past) and among artifacts (a rock vs. the rock used as a paperweight). In such cases, physical identity does not decide overall identity. For minds, as for other sorts of functional devices, physical identity need not decide representational identity nor identity of phenomenal consciousness.

### Acknowledgements

We would like to thank Fred Dretske, Michael Tye, the editor and two referees of this journal for helpful comments on earlier versions of this paper. We also thank the Undergraduate Research Office of the University of Delaware.

### Notes

- [1] While a view like Rey's attempts to treat qualia as similar to narrow content, we will still call this view an internalist view. His is clearly not an externalist view (as are the views of Tye and Dretske) because the qualitative character of sensory states does not derive from the external (broad or wide) properties represented on a view like Rey's. And although the qualitative content of mental states does not derive from the chemical nature of the brain, but from its functional role of various syntactic states on a view like Rey's, it is still true that the qualitative content that there is supervenes on what is inside the head (or skin) (Rey, 1997, p. 303). So this type of theory is clearly closer to an internalist's than to an externalist representational theory. Lastly, Rey's view attempts to "deny the reality of qualia" by locking them within the intentional narrow content of qualia reports (Rey, 1997, pp. 301–308). For these reasons an account like Rey's falls outside the type of squabble we are trying to resolve in this paper. Independently of Rey's view of qualia, we find narrow content views of any sort to be rather implausible, but will not argue that here since we have done so elsewhere (Adams, 1991; Adams *et al.*, 1990) and it is independent of the main themes of this paper.
- [2] Tye assures us that if Swampman's eyes train on his environment causing internal states to co-vary with events in his environment and he produces appropriate behavior accordingly, then he will have conscious states. Tye tells us: "I agree that in general a physical duplicate of a representation need not be a representation; but if the duplicate tracks something in a context in which it is playing the appropriate functional role, then it is a representation." This is consistent with our claim, of course. For then there are internal representations, by his co-variational account of representation (explained below). We challenge this account later in the paper. And as we shall see below, things also depend on whether "functional role" is a teleological notion. Tye seems to want to say no—in which case our objections below apply to his non-teleological account. If he makes functional role teleological, then it transforms into Dretske's account.
- [3] In correspondence, Tye maintains, "I claim that it is conceivable that I came from the swamp ... Any view that denies it strikes me as absurd ... My view has no such consequences ... For a swamp duplicate of me will have experiences." So while Tye sees no absurdity in principle in the possibility that one may discover that one is a swamp duplicate, from the quote he does appear to think it is ridiculous on Dretske's account. We will try to show that it is not ridiculous, even on Dretske's account.

## References

- ADAMS, F. (1979). A goal-state theory of function attribution. *Canadian Journal of Philosophy*, 9, 493–518.
- ADAMS, F. (1991). Causal contents. In B. McLAUGHLIN (Ed.) *Dretske and his critics* (pp. 131–156). Oxford: Basil Blackwell.
- ADAMS, F. (2003). Thoughts and their contents: naturalized semantics. In S. STICH & F. WARFIELD (Eds) *The Blackwell guide to the philosophy of mind* (pp. 143–171). Oxford: Basil Blackwell.
- ADAMS, F., DREBUSHENKO, D., FULLER, G. & STECKER, R. (1990). Narrow content: Fodor's folly. *Mind & Language*, 5, 213–229.
- BLOCK, N. (forthcoming). Mental paint. In M. HAHN & B. RAMBERG (Eds) *Essays on Tyler Burge*. Cambridge, MA: MIT Press.
- DAVIDSON, D. (1987). Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, 60, 441–458.
- DRETSKE, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT/Bradford.
- DRETSKE, F. (1988). *Explaining behavior*. Cambridge, MA: MIT/Bradford.
- DRETSKE, F. (1989). The need to know. In M. CLAY & K. LEHRER (Eds) *Knowledge & skepticism*. Boulder: Westview Press.
- DRETSKE, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT/Bradford.
- DRETSKE, F. (forthcoming a). Knowing what you think. In R. SCHANTZ (Ed.) *The externalist challenge: new studies on cognition and intentionality*. Berlin: de Gruyter.
- DRETSKE, F. (forthcoming b). How do you know you are not a zombie? In B. GERTLER (Ed.) *Privileged access and first-person authority*. London: Ashgate.
- ENC, B. & ADAMS, F. (1998). Functions and goal-directedness. In C. ALLEN, M. BEKOFF & G. LAUDER (Eds) *Nature's purposes*. Cambridge, MA: MIT/Bradford.
- FODOR, J. (1990). *A theory of content and other essays*. Cambridge, MA: MIT/Bradford.
- FODOR, J. (1994). *The elm and the expert*. Cambridge, MA: MIT/Bradford.
- LEVINE, J. (2001). *Purple haze*. Oxford: Oxford University Press.
- REY, G. (1997). *Contemporary philosophy of mind*. Oxford: Basil Blackwell.
- SEARLE, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- TYE, M. (1995). *Ten problems of consciousness: a representational theory of the phenomenal mind*. Cambridge, MA: MIT/Bradford.
- TYE, M. (2001). Phenomenal consciousness: the explanatory gap as a cognitive illusion. *Mind*, 108, 705–725.
- TYE, M. (2002). *Consciousness, color, and content*. Cambridge, MA: MIT/Bradford.