What is Really Wrong With Representationalism?

Søren Harnow Klausen

University of Southern Denmark

Abstract

In this paper I argue that the main problem with representationalism (as defended by Tye and Dretske) is not the view that all mental states are intentional. Contrary to what is commonly assumed, the latter view – “intentionalism” – is neither revolutionary nor particularly problematic, at least when intentionality is understood in a sufficiently broad sense. The real trouble with representationalism is the way it construes the intentionality of phenomenal states. Externalist theories of intentional content are ill suited to capture the qualitative character of such states, which possess a fine-grainedness and significance that is left out by an account in terms of purely objective properties. If this conclusion is accepted but intentionalism maintained, the problem spreads to the theory of intentionality in general, as it becomes necessary to restrict the scope of externalist theories and acknowledge that the intrinsic properties of intentional states play a crucial role in determining their object.

Representationalism in the philosophy of mind is the view that all the characteristic properties of the mental can be explained in terms of representational content. According to the proponents of the view – notably Fred Dretske (1995) and Michael Tye (1995; 2000) – even phenomenal consciousness, which is arguably the most puzzling aspect of the mind, can be fully accounted for on the basis of an externalist theory of representation.

A crucial premise of representationalism is that all mental states are representational or “intentional”, and that their nature is exhausted by their representational character. This thesis, which I shall refer to as “intentionalism”, has often been confused with representationalism. And even though many philosophers do manage to keep the two doctrines apart, the discussion of representationalism has focused almost exclusively on intentionalism, which is widely assumed to be the weakest link in the representationalist chain of reasoning.

I will argue that this assumption is mistaken. Intentionalism is not – or need not be – a particularly provocative doctrine. When considered in isolation, independently of the representationalist superstructure, it turns out to be
acceptable even to philosophers of a fairly orthodox orientation. The real trouble with representationalism is the second step: the attempt to account for the representational character of phenomenal consciousness by means of an externalist theory of intentional content. For such a theory does not seem able to accommodate the kind of content that can plausibly be ascribed to phenomenally conscious states.

Representationalism gets its strong appeal by attempting to combine two approaches to the mental both of which are very attractive: a phenomenological, first-person description of consciousness and an externalist, naturalist theory of representation. Even representationalists have to admit that these approaches are very different and do not seem to be immediately compatible. But they think that once the link between consciousness and intentionality has been established, they will turn out to supplement each other in exemplary fashion. This, I maintain, is far too optimistic. The two approaches remain just as difficult to reconcile as they initially seemed. So if I am right, we might be facing a dilemma: we can adopt a phenomenology-based intentionalism or be externalists about mental content, but we cannot have it both ways.

I

To begin with, let me distinguish some varieties and elements of representationalism. At the root of the theory we have intentionalism. But this is itself a complex view, comprising the following two doctrines:

(I1) All mental states are intentional.

(I2) Every aspect of the mental is intentional.

A common strategy for non-intentionalists is to attack I2, while remaining more or less agnostic with regard to I1 (see e.g. Block 1995, 19f). The affirmation of I1 and the denial of I2 is sometimes called weak intentionalism, whereas their joint affirmation is called strong intentionalism (Crane 2001, 83ff.).

This is good enough for a start, but some finer distinctions can and should be drawn. Intentionality is a very general and to some extent also a vague concept. It can hardly be otherwise as long as one is trying to frame the issue without buying into any particular theory on the subject. According to the
popular definition, an intentional state is one that is “directed at” or “about” an object, and thus has a content. But this is not very informative if it is left open how exactly “object” and “directedness” or “aboutness” should be defined. Intentionalists are prone to widen these notions, claiming, for instance, that the classical sense-data theories were also species of intentionalism, because they held that the mind is directed at an object in acts of perception (see e.g. Crane 2001, 138). And representationalists end up by conceiving of the “directedness” of intentional states as genuinely relational; this clashes with the orthodox position, according to which it is an intrinsic and merely quasi-relational property. So there is room for quite different views within the intentionalist camp, depending on how liberally or narrowly the key terms are defined.

And there is a further complication. Historically, the following view has been very popular: For a mental state to be intentional in an interesting sense, it must have a distinct, relatively “circumscribed” and unitary object, an object which appears as an organised whole and so to speaks stands out from its background. Not all mental states have such an object. Primitive states – or state components – like sensations and feelings obviously do not. But these states nevertheless have an intentional potential or value: they are not intentionally irrelevant or inert, but contribute to the intentional content of the more complex (e.g. perceptual) states of which they may become part. They are necessary ingredients in the content of at least some intentional states, but they are not sufficient for intentionality. For an experience to count as an experience of an object, in a unequivocal sense, the properties that can be said to be represented by the primitive states must be integrated in a way which confers on them certain structural properties, like determinate modes of combination, gestalt-qualities and the like (traditionally, such a process of “objectification” has been called synthesis).

All this indicates that it is not intentionalism per se about which philosophers of different orientation disagree. For it is usually stated as follows:

(I3) All experiences that are alike in their intentional content are alike in their phenomenal character. This is, of course, a doctrine of supervenience. Most contemporary intentionalists take it to be saying that
(14) Phenomenal character supervenes on intentional content.\textsuperscript{5}

But this is not the only consistent interpretation. In fact, (13) could just as well express the reverse relationship:

(15) Intentional content supervenes on phenomenal character.

This is not just an idle thought-experiment. (15) should be taken seriously. For one thing, contemporary intentionalisists often come close to identifying intentional content with phenomenal character. That makes the supervenience relation symmetrical. One can still choose to say that it is the content which determines the phenomenal character, and not the other way round. But this will simply reflect one’s general ontological outlook. The supervenience thesis (13) itself is completely neutral with respect to the order of determination. And I assume that until semantic externalism became popular, most if not all philosophers held the intentional content of a mental state to be determined by the way things seem to the subject, i.e. by the phenomenal character of the state.

However, (15) might not be a completely adequate expression of the traditional view. As I just noted, many philosophers have believed that there are other ingredients of intentionality than phenomenal character, at least when the latter is understood narrowly – i.e. as meaning only sensational quality. So the traditional view should rather be expressed as follows:

(16) Intentional content supervenes on phenomenal character plus certain additional factors.

This still accords with (13), which is assumed to be the essence of intentionalism. Clearly a defender of the traditional view will be less sympathetic to (11) and (especially) (12). But even here there is room for compromise. A traditionalist is likely to insist that primitive phenomenal states could be present without the additional factors needed for the occurrence of genuine intentionality. He might even assume that this possibility is realised in various primitive animals like, say, jellyfish or caterpillars. Presumably, such animals have sensations of hunger, pain and lust; but their experiences remain “unfocused” and do not present the animal with a clearly circumscribed object. This does not contradict the natural assumption that the experiences have an impact on the behaviour of the
animals; the idea is that it influences their behaviour directly, without being mediated by intentional experiences.

The standard intentionalist reply (which goes back to Brentano, the founding father of the position) is that even such a dull and indistinct experience is a form of intentionality. After all, the organisms in question are – ex hypothesi – acquainted with various properties; something is presented to them, even though this “something” is probably of a very blurry nature (Brentano 1924, 127f.).

That can hardly be denied. But it also makes it clear that the question is largely terminological. Intentionalists and non-intentionalists agree on most descriptive issues. They acknowledge that there are huge differences between various types of conscious states, e.g. between seeing that there is a car in front of the house and feeling elated. They also acknowledge that consciousness has an internal complexity, that it always involves both a subject and something which is presented to it, a particular “seeming” or quality or whatever. The non-intentionalists focus on the differences and find them big enough to warrant a distinction between those states that have a genuine object and those that do not. The intentionalists point to the similarities and argue that the differences are only of degree. This is not only a discussion about the use of words, since the parties also disagree about the size and significance of the differences among various state types. Still, it is clearly an exaggeration to say that it is a manifestation of “the greatest chasm in the philosophy of mind” (Block 1995, 19).

By using a more liberal terminology – i.e. by stretching the notion of intentionality – traditionally-minded philosophers should be able to accept I1 and I2. If they widen the notion of phenomenal character, making it cover the “structural” elements in perception as well, they can also accept I5. And once they have widened both the notion of intentional content and that of phenomenal character, they can go on to identify them, which in turn makes the supervenience relation symmetrical. Hence a traditionalist can even accept (I4). So intentionalism is simply not the sticking point in the philosophy of mind.

II

Apart from the terminological confusion, one reason why intentionalism is widely assumed to be an unorthodox view is that it is associated with
representationalism, which embodies a much more special and controversial
type of theory about the nature of intentionality. The case for representationalism is not
nearly as strong as the case for intentionalism. And as I have said, one might
even wonder whether the two views are really compatible.

To get an idea of what I am hinting at, note that representationalists almost
invariably try to motivate their theory by adopting the traditional,
phenomenological perspective on consciousness. Their first – and apparently
chief – objective is to establish the *intentionality of the phenomenal*. This means that
even representationalists accept (I5), at least temporarily. They do as a matter of
fact assume that intentional content is determined by phenomenal character.
Thus we have Alex Byrne saying things like:

> To convey the content of an ordinary visual experience, it is sufficient to
say [e.g.] that it seems to the subject that a red bulgy tomato is on the

And Michael Tye likewise argues for representationalism by examining
experiences “from the inside”, looking for intentional ingredients in their
subjective feel:

> Try to focus your attention on some intrinsic feature of [an] experience
that distinguishes it from other experiences, something other than what
it is an experience *of*. The task seems impossible … The phenomenal
character of your experience … must itself be representational (1995,
136).

The representationalists might reply that this is nothing but a heuristic device:
they begin by describing the pre-theoretical facts and intuitions and then go on
to provide a deeper explanation of them. Tye actually goes further than that. He
suggests that by turning one’s gaze inwards and focusing on the intrinsic features
of one’s experience, representationalism can be vindicated almost immediately:
due to the transparency of experience, one’s awareness “slips through it” to the
“externally instantiated” properties (loc. cit.) and it thus becomes obvious that
“by being aware of the external qualities, [one is] aware of what it is like for
[one]” (2000, 47). In other words, Tye seems to assume that (14) can be
established on the basis of (15).
But this is a little rash. What introspection reveals is merely that an experience has an intentional object (in a broad sense) which does not – at least in most cases – appear to be a mental phenomenon. It does not tell us anything more about the status of this object or how it is individuated. Or rather it does say something about its individuation, but this is not something which supports representationalism: the object is determined by how it seems to the subject (cf. the quotation from Byrne). The intentional content of a mental state is identical to a certain subjective mode of appearance.

But notice that this way of putting it – which the intentionalist cannot but accept – sounds like a textbook formulation of internalism about mental content. Indeed, it is reasonable to assume that the relatively uncontroversial character of intentionalism stems from the fact that it is perfectly compatible with – and even can be said to support – a form of internalism. The traditionalists mentioned above would have no qualms about embracing intentionalism, as long as intentional content is seen as a function of the phenomenal character (together with other internal factors), and neither of these factors are individuated externally. Tradition has it that both phenomenal character and intentional content are intrinsic properties of consciousness. This tallies well with intentionalism. Representationalists, on the other hand, must go (or leap?) from recognising that the content is (or appears to be) determined by intrinsic features of experience to eventually locating it outside the head.

Yet we must be careful not to beg the question against representationalism. For its defenders do not accept the usual contrast between phenomenal character on the one hand and externally individuated content on the other. They are out to externalise the phenomenal as well. Insofar as they succeed in this, the distinction between internalism and externalism will have collapsed. The so-called internal turns out to be external. So although representationalists may be committed to a kind of superficial or prima facie internalism, this does not preclude them from ultimately embracing a strong form of externalism and take this to be the deep truth about phenomenal character and intentionality.

On the other hand, the adoption of the phenomenological (and internalist) stance is not just an innocent, preliminary flirt or heuristic device. Representationalists must respect the phenomenological specifications of the intentional content. This is their explanandum. So the theory of representation which they offer as explanans must capture all the qualities and distinctions that figure in the initial phenomenological description. They are not allowed to
suddenly turn more hard-nosed externalist and ignore or downplay the role of such properties. Whenever it seems plausible that there is – or could be – a qualitative difference between two mental states, representationalists must be able to come up with a plausible external basis for it.

This provides a good opportunity to mount an attack on representationalism. In order to rebut it, one must show that the link between the phenomenal and those external factors on which it is held to depend is too weak to be constitutive. The usual way of showing this would be to envisage thought experiments in which the phenomenal character is held constant while the external features that are supposed to be constitutive of it are varied. Yet many contemporary philosophers, notably those with physicalist leanings, question the legitimacy of this method of “conceptual dissociation” (Tye 1995, 184f.; Dretske 1995, 147ff.). And even though I do not share their scepticism, it is surely advisable to steer clear of complicated methodological and semantic issues as far as possible. An obvious way of doing this is to put forward a more internal (sic) criticism: to examine the externalist theory of representation, seeing whether it can meet the requirements laid down by the phenomenological description of consciousness. Of course, this investigation might still result in a dissociation of the two concepts (viz. phenomenal character and externalist content). And one must still at some points rely on intuitions. But it can be done in a way that is more congenial to representationalism and thus, I hope, also more convincing.

III

Externalist theories of intentionality come in many forms, only some of which are relevant to representationalism. Those which take intentional content to depend on practices of ascription, social institutions or original acts of “grounding”, which are themselves expressions of intentionality, normativity or conscious experience, are clearly not eligible. For representationalism is committed to account for all the characteristic properties of all mental states. And it embodies an uncompromisingly realist view of those states and their content. Whereas the social or “impure” versions of externalism have some plausibility as theories of semantic content, which is both an artefact and a theoretical entity, phenomenal states seem to have a much more robust, natural kind-like ontology.
What remains are those theories which attempt a reduction of intentionality to purely objective relations – relations that obtain independently of human practices, norms and intentions. The obvious candidate is some kind of tracking relation: an intentional state represents the object or state of affairs which it “reliably indicates”, i.e. with which it is causally correlated in an appropriately qualified way. This is precisely the account favoured by Dretske and Tye.

Externalist theories of content face two general problems. They must be able to account for the possibility of misrepresentation. This is equivalent to the traditional requirement that an intentional object need not really exist; an intentional experience can be of or about something non-existent. The other problem concerns the fine-grainedness of representation. It is possible to represent the same object in different ways, and one can be ignorant of the fact that two or more representations are of the same object. This must also be explained, in a way that does no smuggle in elements of unreduced intentionality.

I am not going to rehearse the discussion of these problems. As could be expected, all versions of externalism that are still taken seriously provide answers to them. Here we need only evaluate the answers as applied to the case of phenomenal content. I shall focus on Tye’s theory, since it provides a very instructive example of the internal tension which I take to be characteristic of representationalism. (And it is not because he is an easy prey. On the contrary, his theory seems to me the most sophisticated and plausible version of representationalism, because it takes the first person phenomenological perspective very seriously). Tye favours a causal covariance approach, according to which

\[ S \text{ represents that } P \text{ = df If optimal conditions obtain, } S \text{ is tokened in } x \text{ if and only if } P \text{ and because } P \]

where \( S \) is a state of object \( x \) (1995, 101). This is arguably a rather crude version of externalism. It can deal handily with the problem of misrepresentation: misrepresentation occurs if the circumstances are less than optimal. But this apparently elegant and intuitive solution has its drawbacks. First, it is a consequence of the theory that a state has its intentional content by virtue of having a certain modal property. If the conditions were so-and-so, then it would correctly represent such-and-such object. Now this might not be particularly problematic as long as we are merely dealing with cases of content ascription. And
I see no problem in taking properties of this sort to be real and objective. Dispositional theories of colour are usually classified as realistic, and rightly so, since dispositions can be assumed to exist independently of their being manifested or conceptualised. Accordingly, I find it likely that properties like being justified or being knowledge can be analysed along the same lines. But this is because these properties are theoretical and abstract (just like semantic content). And as I noted before, phenomenal states just do not seem to have this kind of complex ontology. They seem to be more concrete, primitive and “categorical” entities, that is, entities which are identical to their actual manifestations or occurrences. They are something we got right here, whereas “justifiedness” is a property that is conferred on a mental state (which need not even be occurrent or conscious) by its entering into a network of external relations, perhaps including relations to counterfactual circumstances.

It thus makes good sense to say that a mental token represents an external object because it would, under normal conditions, have been caused by this object (or else not have come into existence at all). But it sounds rather strange to say that my pain feels the way it does because under normal conditions, it would have been caused by a certain type of bodily damage (or else not have come into existence at all). Whatever phenomenal character is, it appears to be a property which the mental state (or its “vehicle”) has in a more immediate way.

Secondly, the covariance theory appears to be viciously circular. It makes crucial reference to “optimal conditions”, but leaves them unanalysed. Now if you know in advance what the object of an intentional state is, then of course you can specify the optimal conditions by defining them as those in which the state is actually produced by this (and only this) object. But Tye’s definition is supposed to specify the object in terms of the optimal conditions.

It is perfectly understandable why Tye tends to play fast and loose with the notion of optimal conditions. It reflects his partial commitment to internalism (and, accordingly, his half-hearted endorsement of externalism). Like most of us, he has an intuitive idea about what a particular intentional states is supposed to represent, based solely on its phenomenal character (how things seem to the subject). What I want to suggest is, in other words, that he is secretly relying on a chain of definitions like the following

\[ S \text{ represents that } P = \text{df} \ S \text{ is so that it seems to } x \text{ that } P \]
Optimal conditions of representation = df Conditions in which $S$ (which represents that $P$) is tokened in $\tau$ because $P$ (and only because $P$)

But apart from rendering Tye’s official theory of representation trivial, this has the unwelcome consequence that the notion of seeming to be $P$ is left unanalysed. This is something a traditionalist would happily accept – it is precisely the kind of quality which old-school philosophers take to be basic and indefinable. For example, William Alston has advocated a view – “the Theory of Appearing” – which has many features in common with Tye’s. Like Tye, Alston rejects that some higher-order state or process – be it conceptualisation, belief or the like – must be added to a sensory experience in order to make it the case that an external object is being perceived. He argues that sensory experience in general has the same structure as object perception. And he says of his theory that it “brings the external object into the most fundamental description of the state of consciousness in question” (Alston 1991, 56) But Alston also claims that sensing and perceiving is basically a question of an object’s appearing to one in a certain way.$^6$

According to the Theory of Appearing the notion of X’s appearing to S as so-and-so is fundamental and unanalyzable. … the analysis of the concept of object perception given by the Theory of Appearing is of breathtaking simplicity. For S to perceive X is simply for X to appear to S as so-and-so. That is all there is to it (1991, 55).

Since the representationalist wants to carry out a naturalist reduction of intentionality, he obviously cannot accept this view. And yet Tye seems to be driven towards it. At any rate, he does not show how to break out of the circle of intentional concepts. His official theory refers to optimal conditions, which remains an unreduced intentionnal notion. And this is not the end of the troubles. The problem of fine-grainedness seems particularly acute for a theory like Tye’s. Sensibly enough, he demands of theory of intentionality that it should be able to discriminate between intentional objects – i.e. properties – which are necessarily coextensive (1995, 100). But his own theory does not meet this requirement. Necessarily coextensive properties will covary across counterfactual situations; so the covariance theory does not have the power to discriminate between them.$^7$
All this strongly suggests that the representationalist should opt for a more refined theory of representation. The obvious choice would be a form of “biosemantics”, that is, a “teleological” theory of content which accounts for the normative aspect of intentionality in terms of evolutionary history. A mental state represents the object or feature it was selected for representing; the “normal conditions” for performance of the various representative functions of an organism are likewise determined by the organism’s history (see e.g. Millikan 1993, 86f.). Dretske’s representationalism embodies such a teleological theory.

There are strong reasons for doubting that this approach fares much better than the covariance theory. Though the advocates of biosemantics do seem able to discriminate the intentional objects more finely than those who favour “pure” covariance theories, one gets a suspicion that they are really describing them from a particular point of view which is not the point of view of evolution itself (which presumably has no point of view), but rather the expression of an unreduced “intentional stance”. As Fodor says, “Darwin cares how many flies you eat, but not what description you eat them under” (1990a: 73). It surely seems right to say that frogs represent flies rather than little ambient black things, but again, it is far from obvious that this reflects an objective fact about the content of the frog’s representation, and not just our descriptive preferences.

But serious though they are, I shall set those worries to one side and stick to my plan: the question I shall address is how well a teleological theory can account for phenomenal content – or, in other words, whether the status which it grants phenomenal states (when applied to them) tallies with their phenomenological description.

That the teleological theory has problems in this respect is highlighted by the much-discussed possibility of an accidental duplicate, like the Swampman described by Davidson (1986): a creature that is a molecule-by-molecule duplicate of some actual human being, but is spontaneously generated and thus has no evolutionary history. This is of course a thought experiment which is designed to produce a kind of conceptual dissociation. And so it might seem difficult to draw any substantial conclusions from it without confronting the aforementioned thorny issues concerning the ontological significance of conceptual analysis, imaginability and the like. But I think the case presents a particularly serious challenge to representationalism.

First, it should be noted that most advocates of biosemantics – indeed most externalists – have not presented their view as a theory of conscious intentionality.
Millikan urges that the problem of intentionality should be divorced from the problem of consciousness (1984, 13). Fodor likewise assumes that if there is an internal connection between content and consciousness, “then the problem of intentionality is probably hopeless because the problem of consciousness is probably hopeless” (1990b, 128). And Colin McGinn observes that the teleological theory of content “seems neutral on the question of consciousness” (1988, 296) and fears that the inexplicability of consciousness might prevent us from giving a theory of content, insofar as “the subjective and the semantic are chained to each other” (1988, 298).

Of course these philosophers could be overly pessimistic. But I think they point to a real problem: if there is an internal relation between content and consciousness, then the two factors are interdependent, and this in turn threatens to undermine the whole externalist enterprise. For in that case consciousness is really involved in intentionality, and not just an accidental medium. And then the externalist (in casu teleological) analysis might turn out to be redundant – since the basic representative work is really done by consciousness – or so complicated that nobody is able to actually formulate it.

At any rate, there is the following important difference between the problem which the accidental duplicate causes for orthodox biosemantics and the problem it causes for representationalism: Although the proponents of biosemantics are committed to accept a conclusion which has seemed counterintuitive to many – viz. that the duplicate does not have any intentional states – they are not committed to accept the even more radical conclusion that it is not conscious at all. This is because they do not combine their externalism with intentionalism. Indeed, most externalists seem to suppose that the states of neurological duplicates are “solipsistically type-identical” – that the duplicates are alike with respect to their subjective or experiential states. This means that a duplicate that lacks a history is left with something: a phenomenal (albeit solipsistic, non-intentional) consciousness. He is not a zombie. The representationalist, on the other hand, is forced to conclude that there is simply nothing it is like to be a Swampman.

Dretske has in fact endorsed the latter conclusion: Swampman is a zombie (2000b, 257)! Tye, who is generally much more sensitive to common intuitions and phenomenological evidence, finds it unacceptable. This is the main reason why he prefers a nonteleological tracking theory (1995, 154; 2000, 119). But once again, Tye has a curiously ambivalent – and relaxed – attitude towards his
own theory of representation. For in fact he advocates a mixed view: as long as an organism has an evolutionary history (or other kinds of “design conditions”), this suffices to determine intentional content and thus the phenomenal character of its experiences. In those rare cases where there is no such history, having intentional states is a matter of covariance under optimal conditions (1995, 153; 2000, 121). This enables Tye to cope with both the case of a brain in a vat (of which he would otherwise – according to his covariance theory – have had to say that it could not have any phenomenal states, since its states do not track any external states of affairs) and the Swampman case. Yet such a mixed view seems blatantly ad hoc. It implies that two tokens of the very same experience – say, two qualitatively identical visual experiences of a red rose – can not only be caused, but also constituted, that is made up by properties of a wholly different kind (viz. historical and counterfactual). By advancing such a view, Tye has in fact himself presented a case of conceptual dissociation: he has made it clear that the external circumstances he refers to are merely external to (in the ontological sense, i.e. not constitutive of) phenomenal character. For they can vary while the phenomenal character remains the same.

The ad hocness of Tye’s account makes it easy to think up cases which it cannot handle: it can handle a brain in a vat (by going teleological) and it can handle the case of Swampman (by going counterfactual). But what about the natural extension of these cases: swampman’s brain in a vat?

If our intuitions dictate that both the brain and Swampman have phenomenal states (as Tye explicitly admits), then surely they will also dictate that there is something it is like to be a brain which lacks both a history and states that track external states of affairs. Yet according to Tye’s theory, such a brain is a zombie.

Tye discusses a related objection raised by Ned Block (1998): that of Swampman travelling to Inverted Earth, an imaginary planet on which things have complementary colours to the colours of their counterparts on Earth, but which is similar to it in all other respects. The Inverted Earth scenario alone (without Swampman) gives rise to the following problem for representationalist: if colour-inverting lenses are inserted in your eyes just before you are taken to Inverted Earth, you will not notice any phenomenal difference. But after you have become sufficiently embedded in the new environment, your intentional contents will come to match it: you will represent the sky as yellow, because it is so and your intentional states track it. But you will still experience it as blue. Hence experience and content will have come apart.
The standard representationalist reply to this challenge involves an appeal to teleology: your intentional states will not come to match the new environment, because their content was determined by your evolutionary history. So experience and content stays together (Tye 2000, 119). But if it is the history-less Swampman who travels to Inverted Earth, this reply is not available to the representationalist. As could be expected, Tye resorts to his covariance theory in order to handle this case. But it commits him to saying that the phenomenal character does change when the intentional content comes to match the new environment. A typical way of defending this claim, which Tye himself has been sympathetic to (2000, 126; cf. Lycan 1996), is to rely on an externalist theory of memory states, according to which their content is just as externally determined as that of other mental states – that is, determined by their present environment. This would explain that even though the content – and thus the phenomenal character – changes, the subject will not notice any change, since his memories and his present experiences will stay in harmony.

The problem with this suggestion is, as Tye himself notes, that there must be a transition period during which the intentional states of the traveller become adapted to the new environment and thus acquire a new content. But how does the sky look to the traveller in this period – i.e. before the new counterfactuals are established? Tye sees no other solution to this problem than to accept the original assumption of his opponent and admit that there is, after all, no change in sensory representation with the move from Earth to Inverted Earth, even for an accidental duplicate (2000, 135). This means that Swampman’s sensory system misrepresents colours on Inverted Earth. Tye takes this to be consistent with his covariance theory. Since the operation of Swampman’s sensory system has been interfered with (136ff.), optimal conditions do not obtain.

This reply is unpersuasive, for several reasons: First, it shows how little faith Tye has in its own externalism. He is aware that the external relations – in casu the counterfactuals – are too unstable to really sustain the phenomenal character; they may vary while the phenomenal character remains unaltered. He is not prepared to say that there is no matter of fact about what the phenomenal characters are during the transition period, or that there is no such character at all in the absence of the right system of correlations. His way out is typically ad hoc: if intuition dictates that the phenomenal character remains the same in different circumstances, then some set of external relations must be stipulated to “travel with it” (one could say that Tye rigidifies the intentional content of the
state in question). And Tye eventually resorts to his normality requirement in order to accommodate his non-externalist intuitions: the inverting lenses are a distorting factor. But this move is hampered by his failure to explicate the notion of normality. Swampman has not been designed, and his sensory system arguably has no “initial, natural state” (2000, 136), no built-in standard of “functioning well”. This is not to deny that we have the intuition that Swampman misrepresents the colours in his environment. But it is based on the simple fact that the sky appears blue to him when it is really yellow. It has nothing to do with whether and how his sensory system has been interfered with. If Swampman had materialised as spectrally inverted, we would consider the insertion of (re-) inverting lenses into his eyes an improvement of his sensory system, not because it “restored his initial, natural state” (which of course it did not), but because it established a correspondence between the – independently determined – phenomenal character and the objective properties of his environment.

Secondly, in this and other cases – notably that of afterimages – Tye acknowledges the possibility of a more or less global misrepresentation. A given phenomenal state can have the character it has, even though it in fact never does any positive representational job. It need not even do such a job in any nearby possible world. This is consistent with the covariance theory. But it also highlights the weakness of the – purely modal – tie between phenomenal character and external reality. Just as the possibility of illusion and hallucination led the philosophers of earlier times to recognise a more or less autonomous realm of mental phenomena, the possibility of pervasive misrepresentation still indicates that it must be something else than objective reference relations which determines the character of a phenomenal state (it is thus no coincidence that some of the most refined versions of externalism tend to slide into a form of verificationism, insofar as they make the possibility of successful reference a precondition for intentional content (see Fodor 1990b, 119ff.).

Of course, Swampman’s colour sensations would have tracked colours in some circumstances (if the lenses had not been inserted into his eyes). And surely we are inclined to say that an afterimage would be a veridical experience if there really were coloured objects in the appropriate places. But by saying this, we do not express a belief in any kind of external determination. We take the inverse – internalist – route: the phenomenal character of the state lays down its “conditions of satisfaction”; it determines in what circumstances it would be
veridical and in what circumstances it would not. And there is no reason to assume that the case of Swampman’s colour sensations is any different. Each of them has a representationally significant dispositional property, viz. the property of representing a certain colour in certain circumstances. But it must have this second-order, relational property by virtue of some of its actual properties.

Now what about the case of Swampman’s brain in a vat? Could Tye say that here we have another example of global misrepresentation due to interference with the system’s conditions for proper functioning? This would not be of much help, since he would not be able to identify any set of counterfactuals (or any evolutionary determined function) that could be supposed to “travel with” the brain. Besides, even if he could, it would be wildly implausible to assume that these conditions had anything to do with the actual experiences of the brain, since it could – ex hypothesi – have any experiences whatsoever, depending on the whims of the evil scientist. Tye will no doubt reject the basic premise of the thought experiment, viz. that the brain has its experiences merely by virtue of being in the appropriate neural states. But since there is nothing else that could plausible account for them, he is forced to deny that Swampman’s brain in a vat can have any experiences at all. It is of no avail to assume that the stimulation pattern would mimic some actual covariance pattern, since Tye has amended his covariance theory precisely in order to do away with such cases of “transitive” covariance (S covaries with F, which in turn covaries with G), which would otherwise give rise to representational indeterminacy. So he must bite the bullet and embrace the conclusion that Swampman’s brain in a vat is a zombie.

IV

One of the most obvious shortcomings of the externalist theories of representation is that they do not state the sufficient conditions for phenomenal representation. According to the covariance theory, any state that covaries with some other entity in the appropriate, counterfactual-supporting way can be said to represent it. Tye mentions that tree rings represent tree age (1995, 100), and Dretske uses the example of a speedometer representing the speed of a car (1995, 2). Yet clearly they do not want to suggest that these systems are conscious (it is in fact a little unclear what Dretske thinks, but at least he puts “seems to” and “looks to” etc. in quotation marks when he applies these expressions to gauges or doorbells!). Hence they must bring in some further factor in order to
explain what makes an intentional state into a *phenomenal* representation. For someone who is not already convinced that representationalism is true, this is clearly a ground for suspicion. Could it be that the externalist theory of representation is really marginal or redundant; that the “further factor” is not just something that gives the theory the finishing touch, but the central ingredient in phenomenal consciousness?

The “further factor” which is supposed to yield conscious representation is often mentioned in passing, as if it were something of minor importance or a topic for another, more empirical investigation, perhaps a province of neuroscience (cf. Tye 2000, 63). Externalists speak rather offhandedly about various “sensory modules” (Tye 1995, 138) or “formats” of representation (Tye 1995, 120f.), about “representation consumers” (Millikan 1993b: 126), the importance of being able to store and coordinate representations (Millikan 1993a, 98ff.) or represent oneself as representing things (Dretske 1995, 58) and the like. This is clearly insufficient in the context of representationalism, which stresses the intimate relationship between representation and consciousness. Fortunately, Tye does address – and answer – the question in a more general way: for a representational content to be a phenomenal character, it must be a) abstract, b) nonconceptual and c) “appropriately poised for use by the cognitive system” (1995: 137ff.).

The first condition may be somewhat controversial, since it might be argued that we perceive individual objects and property occurrences rather than types. But I shall not press this point here. The second condition seems unobjectionable, at least to someone who favours a broadly phenomenological or commonsense approach to consciousness (though one might wonder if there are not some representations which are both conceptual and phenomenal. But again, this is of minor importance). The crux of the matter is the third condition. It implies that although phenomenal content is conceived as nonconceptual and nonpropositional, i.e. as not involving any higher-order mental states, it is not wholly independent of such states after all. There can be phenomenal states that are neither conceptualised nor noticed, but an organism which is not capable of having any beliefs at all cannot have any phenomenal states (1995, 144).

So the idea is that representational content becomes conscious by becoming a part of a cognitive system. But what exactly is the embedding in such a system supposed to do to the content? It is surely reasonable to assume that a piece of phenomenal content will “stand ready and in position to make an impact on the
belief/desire system”, (1995, 138) insofar as there is such a system at hand. That is what I meant by saying that every phenomenal state has an intentional “value” or “potential”. But “standing ready and in position” does not sound like the description of an internal, constitutive property. It rather sounds as if Tye implicitly concedes that the content has its character prior to – and thus independently of – its entering into the functional relationship.

At any rate, it might be claimed that Tye eventually succumbs to functionalism: what makes the representational content conscious is its having a certain functional role. It “attaches to the output representations of the relevant sensory modules” (1995, 143) and supplies the inputs for certain cognitive processes (1995, 138). Apart from undermining Tye’s official opposition to functionalism, this confirms the suspicion that the externalist theory of intentionality is redundant. For if the decisive factor is the mediation between the sensory modules and the cognitive processes, what difference does it make whether the mediating item or “vehicle” has an externalist content or not? Is it not sufficient that it occupies the right functional role? This would make it possible to grant that Swampman’s brain in a vat could be phenomenally conscious. And if Tye does not agree that it is the functional role alone which confers consciousness on the representation – which he cannot do without abandoning his representationalist programme – he must at least come up with an explanation of how the combination of externalist content and functional role happens to produce phenomenal consciousness. But it is extremely difficult to imagine what such an explanation might look like.

There are other problems with representationalism which might cause it to slide towards functionalism. It presupposes that every phenomenal state type corresponds uniquely to an external property type, which determines its content. But some phenomenal states may be mere side-effects; they may be caused by other phenomenal states and thus lack an external correlate of their own, even thought they are phenomenally distinct from the states that give rise to them. Consider the case of a masochist being whipped. I take it that Tye is correct in assuming that the pain he experiences is phenomenally similar to that of a non-masochist (1995, 134). The difference lies in the further feelings which the sensation of pain gives rise to in the two cases. What about the masochist’s feeling of lust? It cannot be explained away as a concept-involving state, since it is clearly phenomenal, as Tye himself notes. But neither does it seem to be a further representation of the bodily damage caused by the whipping or of
physical changes in the genitals (which Tye identifies as the intentional object of an orgasm (1995, 118)). That is, I am in fact prepared to say that the lust does contribute to the representation of the bodily damage; it makes the painful bodily damage feel lustful. But this does not support representationalism. On the contrary, it shows that there may be aspects of the intentional content of an experience which cannot be cashed in terms of the objective properties of its object. For the external correlate is the same in both cases.

A somewhat similar case has been described by Ned Block, who suggests that two pains might differ phenomenally though they are identical with respect to intensity, location and other properties which “language can get a handle on” (1995, 37). Unlike Block, I do not think that this undermines intentionalism, as I find it reasonable to assume that the two pains will have different intentional contents (i.e. different “intentional values”. The point is that they could become part of complex experience, in which case they would “colour” it differently). The moral is that they cannot be externalist contents, since they are “thicker” or “richer” than the “naked” physiological correlate of the two states.11

Cases like these indicate that phenomenal content has a fine-grainedness and “aspectual” character analogous to that of propositional attitudes which cannot be captured by the externalist analysis.12 Tye comes close to acknowledging this by describing the intentional object of certain intentional states in a subjective and metaphorical way, using expressions like “being aware of a general sense of buoyancy, of quickened reactions, of somehow being more alive” (2000, 51) or “sensing physical changes in our ‘body landscape’” (1995, 129). And it will not do to point out that differences in phenomenological character are always accompanied by physiological differences. This is a trivial consequence of the widely accepted view that the mind supervenes on the brain. The representationalist must also show that the discriminating physiological factor enters into the content of the state in question. But in the above cases the physiological factor is not represented to the subject. In the masochist case, the physiological element in the representation remains the same, viz. a certain bodily damage.

Tye appeals to something like the James-Lange view of emotions: moods, feelings and emotions are representations of bodily changes; anger is a feeling of rising blood pressure, of the flushing of one’s face, the heaving of one’s chest etc. (1995, 126; cf. James 1950, 449). But this is not sufficient to drive his point home. For it is possible to concede that moods and emotions are always
accompanied by bodily changes, that they are *caused* by these and even represent them, yet still maintain that the states are not exhausted by the bodily component. The issue is not whether moods, feelings and emotions are essentially “bodily” sensations, which they may well be, even on a non-representationalist view. It concerns the nature of sensory representation: is it exhausted by its external content – in casu the “naked” physiological correlate – or does it involve a particular *way* of making this correlate appear to the subject, a certain mode, quality or medium of presentation?

The latter view has much to be said for it, even though one should not exaggerate the strength of the evidence. For the question is empirical, and of the extraordinary tricky sort which can only be decided by armchair empirical psychology (that is, by phenomenological analysis), since it is not merely a matter of reporting whether some easily recognisable phenomenon occurs or not, but of finding the most appropriate description of some very subtle aspects of our mental life, which most people might never have noticed or conceptualised. Still, a pretty strong case can be made for the existence of a qualitative component in experience which goes beyond its objective content. As some critics of representationalism have pointed out (e.g. Crane 2001, 144), the mere fact that we have different sensory modalities indicates that there is more to perceptual content than the properties of the perceived object. Surely it is reasonable to believe that sensations which differ in modality always have different intentional objects (when these are individuated finely enough, i.e. with reference to their represented properties). This is why the plurality of sensory modalities poses no threat to *intentionalism*: seeing and feeling the same object are states with a different intentional content, since they present their objects in different ways. Differences in sensory modality are relevant to intentionality. But it is much less plausible to hold that they are grounded in the *objective* properties of the intentional object. For it appears that these *can* be identical although the sensory modality differs. One might for example both see and feel the same shape or the same relative location.

Another and possibly stronger reason for believing in an additional, qualitative component in sensory experience is the particular *significance* which many sensations have: orgasms are highly pleasant (as Tye concedes; 1995, 118) and impressive (as Block emphasises; 1995, 32), pain is very unpleasant. These qualities, which seem absolutely essential to the states in question, can hardly be said to inhere in the physiological correlates themselves. Tye explains them as a
result of the impact which the states make on the cognitive system. But if the *painfulness* of pain – and the *pleasantness* of orgasms – is conceived as a purely functional relationship, what becomes of the claim that the distinctive features of the mental can all be accounted for in terms of representational content? It is natural to assume that experiences like pain and orgasm elicit strong reactions *because of* their phenomenal character (this picture informs Tye’s own discussion of mental causation (1995, 18ff.)). Yet according to Tye’s official view, it is really the other way round: though the experiences have their phenomenal character by virtue of their representational content, their particular significance is an independent, functionally constituted property.

It thus seems clear that representationalism – at least the version Tye defends – merges with functionalism. Setting to one side the question of whether this compromises it, we can ask in what way functionalism might help to remove the difficulties we have encountered so far. It can handle the masochist case, since it allows that the nature of some mental states depends solely on their relations to other mental states. Hence it does not require that every distinct phenomenal state has a unique external referent. It can handle the case of qualitatively different pains in the same manner, by assuming that they will play different functional roles. And when it comes to the significance of phenomenal states, Tye has already provided the obvious functionalist answer: it is a question of eliciting the right kind of response, either in the form of further phenomenal states or of propositional attitudes.

Note, however, that a functionalist is committed to say that the pain of a masochist and the pain of a non-masochist are phenomenally different after all, since they have different causal roles. Here, pure representationalism is more in keeping with common sense, since it preserves the intuition that the character of simple sensations is robust or “hard-wired”. Their nature is not affected by their giving rise to various further states. Pain is pain, regardless of whether it causes lust or anxiety, or nothing at all. Besides, though functionalism squares with some of the internalist intuitions which are contradicted by representationalism, it does this in a negative and incomplete way, by merely denying that external relations determine content. It does not accommodate the positive internalist intuition that intentional content is determined by the way things seem to the subject – an intuition which may be deemed unreliable in the case of propositional attitudes, but seems deeply entrenched with regard to sensory states. And the properties which it takes to be constitutive of the mental states
are not significantly more robust or obviously relevant to them as the ones posited by representationalism: they are merely dispositional (the identity of a state depends on the other states that it *would* produce), they are not introspectively accessible (cf. Goldman 1993) and they are highly susceptible to “conceptual dissociation”, as exemplified by the problems of spectrum inversion and “absent qualia”. The latter weakness has been duly exploited by representationalists (Tye 1995, 194f.; 2000, 65f.; Dretske 1995, 72).

Hence in my view, the fact that representationalism slides into functionalism does not provide evidence in favour of the latter. By playing the two theories off against each other, one can bring out the inadequacy of “extrinsicalist” theories in general, i.e. of all theories which define mental states relationally. It seems that if we want to preserve the intuitions underlying both representationalism and functionalism, we must adopt a kind of intrinsicalism instead, either of a non-reductive sort, according to which phenomenal character is basic and irreducible, or of a physicalist variety like the one espoused by Block, which takes it to be a matter of the biochemical constitution of the corresponding brain state.

It may be thought that Tye gets himself into trouble because he tries to accommodate too many intuitions in his theory. Would it not be better to do as Dretske and simply opt for one version of externalism – say, a teleological theory – and staunchly deny that creatures which do not meet the requirements laid down by the theory have any phenomenal states? After all, Tye is sceptical about the relevance of imaginability to ontological questions (1995, 183ff.). No matter how sensitive he is to folk psychology and phenomenological observations, he must at some point part company with the folk, since our pre-theoretical intuitions are not fully compatible with representationalism. He must do it in the case of Swampman’s brain in a vat, so why not begin with the case of Swampman? Similarly, why not bite the bullet and submit that having externalist content *is* sufficient for having a phenomenal character, but add that phenomenal consciousness is a matter of degree, and that gauges and doorbells possess it in such a rudimentary form that it seems appropriate to put “seems” into quotation marks when applying it to these devices.

In order to evaluate this strategy, I can no longer stay clear of the issues concerning the relevance of imaginability, conceptual dissociation and the like.
But I shall discuss them in a cautious way, exploiting the representationalists’ own intuitions as far as possible.

Representationalists are prone to emphasise that it has not been shown that phenomenal character does not depend on external factors (Tye 1995, 153; Dretske 1995, 139), and that even though we may think ourselves able to imagine that the external factors and phenomenal character come apart – and though we are indeed able to distinguish them conceptually – this does not show that they could not in fact be identical (Tye 1995, 188ff.). I shall grant all this for the sake of argument: it is at least epistemically possible that there is a constitutive relation of the sort envisaged by the representationalists. But so what? We still need some positive reasons for assuming that the external factors are in fact constitutive. And these reasons had better be very good, since it cannot be denied that representationalism challenges some fairly widespread intuitions.

A typical line of reasoning is the following, which appears to be Dretske’s main argument for representationalism:

1. All aspects of the mental are intentional (strong intentionalism).
2. Our best theory of intentionality is some form of externalism (e.g. a teleological theory).
3. Externalism is the best theory of the intentionality of phenomenal states.
4. Phenomenal character is externalist content.

The problem with this argument is (2): It is not obvious that our best theory of intentionality is some form of externalism, especially when the theory is supposed to cover not only concepts or propositional attitudes, but the intentionality of phenomenal states as well. Dretske himself aptly remarks that one person’s modus ponens is another’s modus tollens (1995, 129). Hence it is tempting to construct an alternative argument along the following lines:

1. All aspects of the mental are intentional.
2. The intentionality of phenomenal states is best accounted for by some version of internalism.
3. Externalism about content should not be applied to the content of phenomenal states.
(4) Representationalism should not be accepted as a theory of mental states in general.

The representationalist may respond to this by making (2) in his argument stronger: the only theory of intentionality available is some form of externalism (Dretske 2000b, 256). But even if we grant him this – if he is right that internalism, especially of the non-functionalist, “intrinsicalist” variety, is a dead-end, an anti-theory which does not explain anything – the argument is not compelling. For it might be taken to indicate that we simply have (at present) no satisfactory theory of intentionality. My own view lies somewhere in between: I do not think that there is anything scientifically disrespectful or obscure about internalist intentionality (or intrinsicalism about the constitutive properties of mental states), and I believe that it can be the object of serious investigations, especially of a descriptive and taxonometrical sort. But I am quite aware – and prepared to accept – that these investigations are likely to leave the fundamental nature of intentionality unexplained (for contemporary defences of the view that intentionality is a primitive and irreducible concept, see Jacquette 1994, 99ff.; Crane 2001).

Tye favours another, more subtle strategy, which has also been adopted by Dretske. He appeals to the transparency (or diaphanousness) of experience (2000, 46f.):

(1) The qualities of phenomenal states do not seem to be qualities of these states themselves, but qualities of their objects.
(2) This provides strong prima facie evidence for representationalism.
(3) It has not been shown that phenomenal character cannot be externalist content.
(4) Representationalism is the best explanation of the intentionality of phenomenal states.

The problem with this is, as I noted above, that Tye makes too much of the transparency. It is correct that the qualities we experience usually appear to be qualities of external objects, and that we may not find any other distinctive qualities when we introspect our phenomenal states. And this can indeed be said to provide prima facie evidence that phenomenal character is nothing but the properties of external things. But this evidence is quickly defeated, since the
possibility of illusion and hallucination precludes us from straightforwardly identifying the qualities of our experience with the properties of external objects.

Representationalists are of course aware of this. Accordingly, Tye identifies the phenomenal character with the representational content instead. This is not the only option, but I think it is the right one – most orthodox philosophers of mind would agree, especially those who do not want to posit sense data and the like. The problem is that it is left completely open how the content is to be analysed. One can opt for externalism, but this is not supported by the transparency observation. True enough, the experienced qualities seem external. But this is not the kind of externality which characterises externalist content. It is simply a particular mode of appearance (see Fales 1995, 106f.). Introspection does not reveal any counterfactual-supporting covariance relations or historical facts about one’s ancestors. Note that I do not take this to show that representationalism cannot be correct, only that it cannot be based on any introspective evidence. (One could, however, press the point that the qualities we experience seem to be categorical properties of objects, and that they appear markedly different from the way content is supposed to be, on almost any theory of intentionality. For content is widely assumed to be just as transparent as experience; it is that by virtue of which an intentional state represents whatever it does, but it is not literally the object of the state. This is so even according to those externalist theories which claim that the object “enters into” and determines the content; the latter is a more abstract and complex – relational – entity, which can only be revealed in an act of reflection. But I shall resist the temptation to pursue this point. Still, I cannot keep from observing that Tye apparently thinks that brain-state identity theories can be refuted by appeal to the phenomenological facts: “Peer as hard as you like at the neurons … you will not find any phenomenology” (1995, 162). Do we find any phenomenology by peering at the covariance relations instead?).

I have granted that conceptual dissociation may be an unreliable guide to metaphysical possibility. And I shall even grant that observations concerning the basis of self-attribution of mental states do not provide any conclusive evidence against representationalism and other forms of “extrinsicalism” (as argued by Goldman 1993). But we still need a reason to believe that phenomenal character is externalist content. If the representationalists are right about the limitations of traditional philosophical methodology, phenomenal character could turn out to be virtually anything. That may be a problem for their own theory, since it
becomes unclear why the relationship between phenomenal character and externalist content should be an *identity* and not just a counterfactual-supporting correlation.

Now clearly the representationalists’ paradigm is reductive explanations in the natural sciences: they consider the claim that phenomenal character is externalist content an a posteriori necessity of the Kripkean sort, analogous to the claim that water is $H_2O$. But even if we ignore Kripke’s own scruples about extending his view of scientific identification to the relationship between mental and physical states (1980, 151) and similar considerations, there remains an important difference. In the ordinary scientific cases, we are always told a detailed story about how the higher-level properties referred to by our folk-concepts are generated by more basic physical properties. Tye himself rightly demands of a physicalist theory of mind that it should do precisely that: describe a mechanism which provides an explanatory link between the phenomenal and the physical (1995, 17). But it seems that he and other representationalists eventually shirk this requirement. What they provide is nothing but *coextension intuitions*: they appeal to the natural (realist and non-sceptical) assumption that a certain phenomenal character is likely to be reliably correlated with the external state of affairs which it represents. And they further assume that if we have a sufficiently tight – counterfactual-supporting – correlation, we are likely to have a case of identity or at least strong metaphysical dependency as well.

There are at least two serious problems with this reasoning. First, it overlooks that the coextension intuition – or fact, if we do have sufficient evidence to be entitled to call it so – supports the traditional internalist view just as strongly: if the phenomenal character of a sensory state determines its object, and if sense perception is a reliable source of information about the environment, then it will be correlated with the corresponding external state of affairs in precisely the same counterfactual-supporting way. (It should be noted that the internalist can also avail herself of a reliabilist theory of knowledge. There is no necessary connection between semantic and epistemological externalism). It seems that semantic externalists are generally far too confident that the obtaining of the appropriate set of objective correlations will count as evidence for their – and only their – view. As Amir Horowitz puts it, they erroneously assume that “if there is an externalist way of describing a situation, the situation exemplifies externalism” (2001: 331).
Secondly, no matter how robust and pervasive the correlations between the phenomenal states and the external states of affairs turn out to be, showing this still falls short of establishing their identity (see Kim 1998, 5; 11ff.). This would require additional mechanism-descriptions of the sort provided by the paradigm scientific explanations. And the representationalists have not come up with anything like that. They have merely struggled to “carve out” the external factors in a way that is extensionally adequate. Whether they have succeeded − or can succeed − in this is an open question. But it is only of marginal relevance. Even if the representationalists were able to provide externalist descriptions which as a matter of fact specified the content of every phenomenal state uniquely, we could still have reason to doubt whether this specification says anything about the nature of such states. For it is quite common to identify entities by their relational properties without thereby assuming that these properties are defining or constitutive (see Levine 2001, 98). It is reasonable to assume that a certain type of entity is the only one which, as a matter of fact, does a certain job, without thinking that its whole being consists in its doing this job. This possibility can, to repeat, only be ruled out by showing how the relational properties make up the entity.

This point might also serve to bring out how my view differs from that of Colin McGinn, to which it otherwise bears an obvious similarity. McGinn argues that because phenomenal character and intentionality are “chained to each other”, and the former defies explanation, we have reason to fear that the mystery of consciousness will spread to intentionality (1988, 297f.). Yet McGinn ends on a more optimistic note: though we may not be able to fathom the nature of intentionality, the individuation of intentional contents is still open to us (302). Though I do agree with McGinn that the externalist theories of content contain valuable insights and that the intertwinement of consciousness and intentionality does not preclude us from saying something illuminating about the latter, I do not share his diagnosis. For I doubt that the externalist theories can contribute to the individuation conditions of mental states (304). If they could, we would have reason to be more optimistic than McGinn, since stating the individuation conditions of something (in a complete and nontrivial way) and describing its nature are arguably one and the same thing.

We can indeed state the difference between intentional states by describing how their objects differ: seeing a red apple is different from seeing a brown book. But is this a naturalistic account of the “specific content” of each state?
I do not think so. For it only seems unobjectionable as long as it is not subjected to further externalist analysis. When the contents of “red apple” and “brown book” are cashed in terms of causal or counterfactual or historical relations to objectively red apples and objectively brown books, it becomes questionable whether the two states can really be individuated in this way. To support the latter claim, we need the further arguments for identifying phenomenal character and externalist content which cannot be found in the representationalist tracts (and which McGinn doubts can be given at all). Of course, if “individuation” merely means “identification”, then I concur with McGinn’s proposal. We can exploit the stable correlations to make symptom-based identifications, just as we do in countless other cases. This is arguably the way we ascribe mental states to others. But just as behavioural manifestations are not, in these post-behaviouristic times, assumed to be anything more than reliable clues to the underlying mental reality, there is no need to think that the objective properties of objects perceived by ourselves and others are genuine individuation conditions of the content of the corresponding states (except for in the trivial sense that insofar as the states are veridical, and the properties thus rightly represented by the states, the equivalence between the internal and the external description will ensure the adequacy of the latter). Here, as in many other cases, it is important to distinguish between the pragmatic constraints of property ascription and the ontological conditions of property possession. Externalist theories of intentionality may prove successful in stating the former, but we have currently no reason to be optimistic about their ability to capture the latter.

VI

I believe to have shown that the alliance between intentionalism and externalism is more problematic than is usually assumed. I do not pretend to have refuted representationalism; indeed, it is hard to see how this could be done, granted that almost any view on the matter could, in principle, turn out to be correct, that intuitive and introspective knowledge is limited and fallible and so forth. Still, I feel entitled to conclude that as long as no illuminating story about the generation of phenomenal consciousness is forthcoming, it has little to be said for it. And the alleged discrepancy between the rich and multifariously significant phenomenal character of many states and their one-dimensional
externalist content is a strong reason for doubting that representationalism can provide an adequate description of its subject matter.

In order to play the game, representationalists must confine themselves to quite austere descriptions of the intentional objects of phenomenal states; but it is doubtful whether such descriptions can match the phenomenological or folk-psychological ones. The only way to eschew this problem seems to be a slide into eliminativism, i.e. taking the first person perspective much less seriously than originally announced and thus cutting the concept of phenomenal character loose from its phenomenological and folk-psychological moorings. But the appeal of representationalism stems largely from its promise to preserve a large number of folk-psychological beliefs and philosophical intuitions. If it is simply stipulated that phenomenal character is whatever happens to meet the externalist requirements, one has good reason to think that the problem of phenomenal consciousness has been evaded rather than solved.

Where does this leave us? The following options suggest themselves. One can (a) abandon intentionalism and maintain an unrestricted externalism about intentionality, (b) adopt a “mixed” theory of intentionality, according to which the content of some states is externalist while others – notably the phenomenal states – have internalist content as well, or (c) exploit the failure of representationalism by mounting general attack on externalism, reinstating internalism across the board. I believe that (a) would be very difficult to defend, as it seems to me that representationalists (and orthodox intentionalists) have argued very convincingly for intentionalism, at least in its more liberal versions (e.g. I3). What may at a first glance appear to be serious objections to it are only pertinent to representationalism; they bring out the implausibility of an externalist construal of the intentionality of phenomenal states. Though I have some sympathy for (c), I think the case for externalism about at least some kinds mental and semantic content has been made out strongly enough to prevent us from discarding it completely. So I think one should settle for (b). Still, a full appreciation of intentionalism together with a rejection of representationalism may take one pretty far in the direction of (c), since phenomenal character arguably plays a role in determining the content of quite many types of mental states, including propositional attitudes (see Goldman 1993, 123). Externalism may only be salvaged as a theory of belief ascription and the conventional and historical aspects of linguistic meaning (see Crane 2001, 117ff. for a defence of this view).
To many, this conclusion will seem completely unacceptable. But I do not see how it can be avoided. Besides, things are not as bad as they might seem. If we are prepared to maintain that phenomenal character is an intrinsic quality of mental states, which is realised directly by biochemical properties of the brain – and not mediated by functional or representational properties – is it then much less plausible to assume that the same should hold for intentionality, at least in the consciousness-infected, internalist sense? Surely we have virtually no idea of how this property is realised. But neither do we have any idea of how it could be realised by externalist content. What we have is merely an extensive system of correlations, together with general metaphysical and methodological reasons for preferring materialism to dualism. If this gives us sufficient justification for the belief that phenomenal states are realised by physical states, let us rest content with it and not pass off theories like representationalism as illuminating descriptions of the mechanism that is responsible for the realisation.¹³ Let us stick to this modest, realistic view and apply it to the intentionality of phenomenal states as well. And let us then appreciate the externalist theories of linguistic meaning, attitude ascription, information processing, indication and the like, bearing in mind that they are answers to different problems, but valuing them nonetheless.¹⁴

Notes

¹ There are other prominent defenders of representationalism, for instance Harman (1990) and Lycan (1996). But in order not to get myself entangled in exegetical question, I shall not discuss their views, which differ from those of Dretske and Tye in various respects. Still, most of my arguments apply to them as well.

² A partial exception – of which I have only recently become aware – is Levine 2001, 117ff.

³ Versions of this view has been defended by Kant, Frege (1986, 51), Husserl (1984, 80), C. I. Lewis (1929, 59; 131) and H. H. Price (1953, 41f.). The additional element need not be conceptual or propositional, though it is taken to be so in the Kantian tradition. Husserl and other phenomenologists believe in the possibility of simple, pre-conceptual seeing (as described by Dretske 1979), but maintain that there is a difference between having certain sensations and perceiving an object. They think that the latter requires a pre-conceptual act of synthesis or interpretation (see Mulligan 1995, 206f.)

⁴ Tye 2000, 45; Byrne 2001, 204

⁵ Byrne 2001, 204

⁶ Crane (2001, 17) defends a similar view.

⁷ Tye has recently tried to improve his covariance theory by adopting a proposal from Fodor. He suggests that the theory should be supplemented with a further asymmetric dependence condition. The idea is that for S to represent F not only must S causally covary with F under optimal conditions but it must also be the case hat if there is some other feature G such that F
covaries with $G$ under optimal conditions, then were $F$ to fail to covary with $G$, the causal covariance link between $S$ and $F$ under optimal conditions would still hold but that between $S$ and $G$ would be broken (2000, 140). This does suffice to narrow down the class of possible intentional objects. Still, the theory cannot discriminate between metaphysically (as opposed to merely nomologically) coextensive properties.

8 I am indebted to Oliver Bott here.
9 See footnote 7
10 I am indebted to Cathrin Misselhorn here.
12 DeBellis (1991, 314f.) argues that one needs to posit fine-grained phenomenal content in order to account for certain features of musical experience.
13 Note that I do not deny that there is such mechanism (as Searle apparently does (1992, 105)) or claim that we are fundamentally precluded from gaining any insight into it (as McGinn assumes (1988)). All I am saying is that representationalism – in its current state – does not give us any such insight.
14 They are, for instance, of obvious epistemological value, since they support (and are supported by) a realistic metaphysics and a relatively optimistic view of human cognition.

References

Davidson, D.: 1986, “Knowing One’s Own Mind”, *Proceedings and Addresses of the American Philosophical Association* 60, 441-58
Dretske, F., 2000b, “Norms, History and the Constitution of the Mental”, in
Dretske 2000a, 242-258
Content and Other Essays. Cambridge, Mass.: MIT Press, 51-87
Content and Other Essays. Cambridge, Mass.: MIT Press 89-136
Götingen: Vandenhoeck, 30-53
Consciousness and Cognition 2, 364-382 (reprinted in Block et al. 1997, 111-125)
4 (ed. J. Tomberlin), Northridge, CA: Ridgeview, 31-52 (reprinted in Block
et al. 1997, 663-675
Husserl, E.: 1984, Logical Investigations II/1 (Husserliana XIX/1). The Hague:
Martinus Nijhoff
University Press
Lewis, C. I.: 1929, Mind and the World Order. New York: Dover
1997).
Millikan, R. G.: 1984, Language, Thought and Other Biological Categories. Cambridge,
Mass.: MIT Press
Millikan, R. G.: 1993a, “Biosemantics”, in White Queen Psychology and Other Essays
Philosophy 86 (1989), 281-297; page references to 1993)
Millikan, R.G. 1993b. “Compare and Contrast Dretske, Fodor and Millikan on
Teleosemantics”, in White Queen Psychology, 123-133
Cambridge Companion to Husserl, Cambridge: Cambridge University Press, 168-
238
