

Affect: Representationalists' Headache

Murat Aydede (UBC) & Matthew Fulkerson (UCSD)

Vision tells us that the strawberry is spotty red, round, and located a short distance from our hands. While even here controversies abound, there is a plausible and well-supported general account of how vision performs this task: our visual system extracts information about the sensible visual features in the environment, and makes this information consciously available to a subject for further processing and action-preparedness (both motor and epistemic). On this model, simplifying quite a bit, perceptual experience *represents* features of the environment. One of the main projects of recent philosophy of perception and cognitive science has been to fill in the details of this general account. At present there are many options from which to choose, and no clear winner is on the table.

It is difficult enough to work out the details of such an account for even the most basic sensible features, like colors and shapes. And perhaps this explains why so much effort has been devoted to these features in recent years. This focus is fine so far as it goes, but it can have the unintended consequence of obscuring the full richness and variety of perceptual experience. Perception, after all, not only provides information about objects in the environment but also often seems to pass a kind of affective “judgment” on them. This judgment manifests itself in the pleasant and unpleasant character so common to many of our perceptual encounters. There is something especially pleasant and seemingly good about the taste of a juicy, perfectly sweet strawberry on a warm summer day. And there is something especially unpleasant and seemingly bad about encountering an awful smell, like rotting fish, in a confined space.

In both cases, as in the paradigm case of seeing that the strawberry is red and round, there seems to be something — the strawberry — presented to us being a certain way, as *pleasant* or *unpleasant*. But whereas it is initially plausible that there is some objective sensible property — call it *redness* — that our visual system detects and represents, it's not at all clear that there is an objective analogue in the case of the pleasant and unpleasant. Accounting for these features poses a major challenge to any extent account of perceptual experience. Indeed, we believe, and shall argue in what follows, that these features prove to be fatal to a popular class of views known as strong representationalism. This fatality stems from a crucial ambiguity in the nature of affective experience, one that no version of strong representationalist can adequately explain in its preferred terms. We will lay out the details of this argument in stages, starting with a minimal framework for thinking about perceptual affect, followed by a characterization of our main target.

1 Locus of affect: A framework

It is clear that we find many things pleasant or unpleasant. A walk on the beach on a warm day is quite pleasant. Grading a stack of exams on a warm day, however, is not. These are examples of *secondary affect*: cases in which some activity or general state of affairs is deemed in positive instances to be enjoyable or fun, and in negative instances to be unenjoyable or a chore. This is a rough characterization, to be sure, but we mention it only to set it aside in this essay. Such cases are not our intended focus. Our interest lies in a more specific occurrence, in cases of *primary affect*, in which a token sensory experience seems to present us with the pleasant and unpleasant character of things.¹

We start by noting a curious ambiguity found in attributions of primary affect. Let's call the qualities like the awfulness of smells, the pleasantness of tastes, the painfulness of pains, and so on, *affective qualities*.² The ambiguity arises when we ask: what do these qualities qualify in the first instance? It seems obvious that, depending on context, we attribute these qualities both to our perceptual experiences and to their objects. Although it seems to apply in all sensory modalities, the ambiguity is especially robust in the case of smells, tastes, and bodily sensations. Indeed, when we talk about how pleasant the taste of a strawberry is, we may be attributing the pleasantness to our subjective *experience*, or we may be attributing it to whatever objective sensible qualities of the strawberry are responsible for our experience.³ We will assume that

¹ 'Affect' as used here is sometimes called in the literature 'valence' or 'hedonic tone' that can be positive or negative. Hereafter, unless otherwise noted, we shall use 'affect' to refer exclusively to primary affect. Primary affect, as we shall sometimes say, is *perceptual affect*, affect directly presented to us in perceptual experiences. Intuitively, perceptual experiences, when they have an affective phenomenology, seem to present to us the affective qualities of perceptual objects. For this reason, it seems proper to say that the phenomenology of primary affect is the phenomenology of *affect-presenting* experiences, whereas the phenomenology of secondary affect may be identified with the phenomenology of *affect-causing* perceptual experiences (such as the extreme unpleasantness of feeling a lump under my arm after five years of remission of a lymph cancer). We don't claim that the distinction between primary and secondary affect is simple, clean, or clear-cut.

² Throughout the paper we will assume *phenomenal realism* about affective qualities, according to which there is a distinctive episodic phenomenology to experiencing affective qualities when we perceive or introspect them. The phenomenology of affect is a fairly interesting question that needs more serious discussion. One of the main worries is that experiencing affective qualities doesn't have the same kind of phenomenology as that of experiencing standard sensible qualities (like colors for instance). Characterizing the difference turns out to be difficult and controversial (see e.g., Robinson 2006; Clark 2005; Aydede 2000). But for our critical purposes in this paper, we don't need to address this issue: what we mean by 'phenomenal realism' is that there is some *episodic* affective aspect to our experiences that makes an introspectible difference. Affective qualities, then, are just those features (whatever they are) that are partly or wholly constitutive of these aspects. They may not be intrinsic or simple, they may not even be qualitative in the sense of there being qualia for which a quality space can be specified through standard multidimensional scaling experiments (Clark 1993, 2000, 2005). In this weak sense, phenomenal realism about affect is common ground between us and our opponents. However, some early representationalists such as Armstrong (1968) and Pitcher (1970), who advocated affect as purely attitudinal responses to sensation, could be interpreted as phenomenal eliminativists about affective qualities.

³ In what follows, we will assume (for convenience) a primary quality view of all sensible qualities including so-called secondary qualities. On such a view, all sensible qualities are extra-mental objective/physical properties (or relations). But nothing very important hangs on this for our present purposes: all our opponents assume such a view of sensible qualities.

both kinds of attributions can be correct. So suppose, in the case where one takes a good bite from the strawberry, the following express two judgments that are true:

- (1) This strawberry is pleasant.
- (2) My (taste) experience of it is pleasant.

(1) expresses a *de re* judgment attributing a positive affective quality, *pleasantness*, to the object of my perceptual experience (or, perhaps more precisely, to the complex of the strawberry's sensible qualities objectively understood — we will mostly ignore this for convenience). Hence it is what we will call a *first-order* or *perceptual* judgment directly (i.e., non-inferentially) prompted by the experience of the strawberry. (2), on the other hand, is an *introspective* judgment attributing the affective quality directly to the *experience* of the strawberry (rather than to the strawberry itself).

Can (1) and (2) attribute the same affective quality? Can strawberries and their experiences be pleasant in exactly the same way or sense? The intuitive — and, as far as we can tell, widely shared — answer is no. The qualities attributed in (1) and (2) are distinct properties. Is there some sort of a theoretically interesting connection between the two properties? Here, again, the intuitive answer is yes, there is. Whatever they are, these properties are not just accidentally related. It is plausible to assume that there is some kind of dependency relation between them.⁴ After all, the instantiation of one is somehow implicated in the instantiation of the other. Putting aside the issue of what these properties are and the exact nature of this dependency, for the moment, we can now define two incompatible views about the direction of the dependency:

THE EXPERIENCE VIEW (EV):

- Affective qualities *primarily/fundamentally* qualify perceptual experiences rather than their objects, and
- For any x and affective quality A , if x is A and x is *not* an (aspect of an) experience, then x is A only *derivatively* (i.e., x is A only because its experience is A).

THE OBJECT VIEW (OV):

- Affective qualities *primarily/fundamentally* qualify the objects of perceptual experiences rather than the experiences themselves, and

⁴ We will see that both these claims (that the properties are distinct and that there is dependency relations between them) are common ground between us and our opponents.

- For any x and affective quality A , if x is A and x is an (aspect of an) experience, then x is A only *derivatively* (i.e., x is A only because its object is A).

Note that, with some fairly natural assumptions, these views entail another dependency ordering among our ways of epistemically accessing affective qualities. EV entails that introspective access to affective qualities is primary/fundamental. On this view, if we ever *perceive* affective qualities or make (first-order) perceptual judgments about them, it is only because we can *introspect* those that qualify the experiences. Call this the INTROSPECTION-FIRST VIEW (IFV). On the other hand, OV entails that the *perception* of affective qualities is primary/fundamental, and our ability to *introspect* them (if we ever do) or make introspective judgments about them is derivative — i.e., depends on our ability to perceive them or make perceptual judgments about them. Call this the PERCEPTION-FIRST VIEW (PFV). We will return to this distinction later.

The picture we have drawn so far is fairly abstract and somewhat formal. It says nothing about what affective qualities are or what the dependency relation among them is. It does not assume any particular view about how introspection or perception works, or what it takes to make perceptual or introspective judgments. It is neutral on the nature of perceptual (and, for that matter, affective) experiences. It does not say whether EV or OV is correct. It takes a cue from the way we attribute affective qualities and sets up a fairly broad framework with very minimal constraints in it. We are inclined to believe that pretty much anyone who is a phenomenal realist about affect can agree with this framework.

However, we will show that this framework has enough meat to cause serious headache for the defenders of (strong) representationalism about phenomenal character. In particular, the framework presents a fatal dilemma for defenders of strong representationalism. By their own lights, they should prefer to adopt a version of OV. But as we will show in some detail, OV is a thoroughly implausible and deeply problematic account of affective qualities and cannot be adopted by the strong representationalist. So that leaves some version of EV as the only plausible account of affective qualities, but EV is inherently anathema to both the spirit and the letter of strong representationalism. We show that on both standard and extended versions of the view, EV and strong representationalism cannot be plausibly combined. Given the minimal framework, this leaves the strong representationalist with no way to account for affective qualities and, given the ubiquity and centrality of affective qualities in perceptual experience, this poses a serious problem for the view.

Before we spell out our argument in more detail, we need to make clear the kind of representationalism we will target in this paper.

2 Representationalism and the problem of affect

Intentionalism in philosophy of perception is the view that all perceptual experiences are representational. Intentionalist theories vary widely. They can be reductionist or non-reductionist, naturalist or non-naturalist, pure or impure, wide or narrow, Russellian or Fregean,

restricted or non-restricted. Some of these classifications cross-classify the theories, and are not completely independent of the others. Furthermore, many of these distinctions are known under different names in the literature, sometimes with (slightly) different extensions. So to forestall potential confusion we start by clearly stating our target: We are interested in the group of intentionalist theories that are reductionist, naturalist, pure, wide, Russellian, and non-restricted, and sometimes known as *strong representationalism* (hereafter, SR). Defenders include Harman (1990), Dretske (1981, 1995), Tye (1995, 2000), Byrne and Hilbert (1998, 2003), Byrne and Tye (2006).

According to SR, the entire phenomenal character of an experience is one and the same with, or is *entirely* determined by, the wide representational content of the experience. The semantic externalism assumed by representationalism thus yields phenomenal externalism that says that the phenomenal contents of the experiences of physical duplicates may differ depending on their environments.

Representationalists defend or assume naturalistic accounts intended to explain how the vehicles of experiences realized in the brain acquire their representational content.⁵ Furthermore, they propose or assume naturalistic constraints on the vehicles (or their content) that are intended to explain why only experiential representation has an essential and distinctive phenomenology in the way it does.⁶ The representationalist project is therefore a metaphysical project: the ultimate aim is to understand conscious experiential phenomenology completely in naturalistic and reductionist terms.

Given that conscious phenomenology is essentially introspectable (at least for creatures of a certain sort), representationalism entails the STRONG TRANSPARENCY (ST) thesis:

⁵ For example, Dretske writes in the prologue to his 1995 that his Representation Thesis (a version of SR) holds two tenets: “(1) All mental facts are representational facts, and (2) All representational facts are facts about informational functions.” He concludes that this view provides “a satisfying account of the qualitative, the first-person, aspect of our sensory and affective life—distinguishing in naturalistic terms between *what* we experience (reality) and how we experience it (appearance).” Later in his lectures, he fills in the details of this naturalistic account: “According to the Representational Thesis, the facts that make what is in the head mental, the facts that convert electrical and chemical activity in the cortex into blue-dog experiences, are facts that are not identifiable by looking, exclusively, at what is in the head. What makes a certain pattern of electrical activity in the cortex into a blue-dog experience is a fact about what this activity represents, what it has the function of indicating” (1995: 36-37). Dretske thus defends a teleological form of indicator psychosemantics. For a summary of varieties of psychosemantics proposed for conscious sensory representations, see Lycan (2008).

⁶ As Tye writes, “The broad picture I have here of perceptual sensations draws a sharp distinction between these states and beliefs or other conceptual states. Perceptual sensations, in my view, form the outputs of specialized sensory modules and stand ready to produce conceptual responses via the action of higher-level cognitive processing of one sort or another. So perceptual sensations feed into the conceptual system, without themselves being a part of that system. They are nondoxastic or nonconceptual states. This, I want to stress, does not mean that perceptual sensations are not symbolic states. But, in my view, they are symbolic states very different from beliefs” (Tye 1995: 103-104). Tye (1995, 2000), like Dretske (1981, 1995), puts further conditions on a representation to count as sensory such as being poised to be consumed by central cognitive systems, abstract, non-object-involving, etc.

(ST) Any quality that we (can) epistemically encounter in the introspection of an experience is a quality only (widely) *represented* by this experience, thus not a quality of the experience (*a fortiori*, not an *intrinsic* quality of the experience).⁷

Because some version of ST is intuitively deemed self-evident on ordinary phenomenological grounds, many representationalists argue that this entailment is an argument in favor of their representationalism.⁸ So, if, in having an experience, there are any qualities we can introspect that are not attributable to the objects of the experience, representationalism is refuted. This is not something we argue for here, but which is assumed on all extant SR accounts. We intend to show that affective qualities are such qualities. We will show that representationalism is false on account of the fact that it cannot explain affective phenomenology in its preferred terms.

To the extent that the affective aspects of experiences are phenomenal, representationalists are committed to giving a representationalist account of them on a par with accounts they give for other sensory-discriminative aspects of experiences. Yet very few representationalists have taken up the challenge and attempted to develop a detailed account. Similar worries arise for other phenomenal experiences such as the so-called intransitive bodily sensations such as pains, itches, tickles, orgasms, etc.,⁹ as well as for general bodily feelings, moods, and aspects of emotions. Many of these pose a challenge precisely because they don't seem representational at all. The worry about affect is different, since it arises even if we allow that affective qualities are represented (perhaps derivatively) in perceptual experience. So they pose a novel and especially difficult challenge to SR, and in general representationalists have been coy in directly tackling them.

Michael Tye has been a remarkable exception to this trend among representationalists.¹⁰ Since around mid-1990's until very recently, he has written extensively — significantly more

⁷ The notion of 'epistemic encounter' is left deliberately vague in this formulation in order to avoid making substantive assumptions about how introspection works and what it requires. Nevertheless, we will be assuming that introspection, when it yields introspective *knowledge* as it often does, requires appropriate *concepts*. This should not be controversial between us and our opponents in this essay — see below for more on this.

⁸ See especially Harman (1990) and Tye (2000, 2002) — cf. Dretske (1999: 103): "(1) Conscious perceptual experiences exist inside a person (probably somewhere in the brain). (2) Nothing existing inside a person has (or needs to have) the properties one is aware in having these experiences." However, we don't want to claim that the way the transparency intuitions are supposed to support SR is very clear in the writings of these authors. In fact, despite the fast growing literature on the topic, we don't think there is any emerging consensus about what exactly transparency is and what it is supposed to show about theories in philosophy of perception. See Aydede (forthcoming) for an attempt to clarify the thesis in more detail. Accordingly, our point in the main text can be taken as a way of clarifying our target of criticism: roughly, we want to criticize the kind of representationalism that has ST as its consequence. We don't think this should be controversial — whatever the more accurate historical or philosophical scholarship on existing literature turns out to reveal.

⁹ For in-depth discussion of these difficulties, see Armstrong (1962, 1968), Aydede (2001, 2009, 2010).

¹⁰ Tye (1995, 1997, 2000, 2006a, 2006b). For other representationalist treatments of pains and other bodily sensations, see Harman (1990), Dretske (1995, 1999), Byrne (2001), Seager (2002), Bain (2003, 2007); for the *affective* aspects of pains, see especially Bain (2013) and O'Sullivan and Schroer (2012). Our criticism below will

than any other representationalists at any rate — on how to extend representationalism to treat these problematic cases. And more recently, in a work co-authored with Brian Cutter (Cutter & Tye 2011), Tye developed and defended a sustained representationalist account exclusively of experiential *affect*. Although their account is specifically directed towards the negative affective character of pain experiences (the *painfulness*), they are explicit that the account is meant to generalize to the positive as well as negative affective phenomenology of *all* experiences. We believe the account C&T develop is paradigmatic for representationalism. Indeed it is almost a textbook example of what a principled (strong) representationalist would or could say about *affective* phenomenology. Therefore, our criticism, although it takes some aspects of C&T's account as its stalking horse, is intended to generalize to the core thesis of representationalism itself; that is, roughly to the intentionalist views that are reductionist, naturalist, pure, Russellian, wide, and non-restricted. Such views say that all experiential phenomenal content (non-restricted) reduces without remainder (reductionist and pure) to wide (Russellian) representational content of a certain sort, of the sort that completely admits of a naturalistic psychosemantics (some version of indicator and/or consumer semantics as applied to mental representations that play a certain kind of role in the informational architecture of creatures capable of conscious experiences).

While these various choice points may make it seem like we are picking out a narrow target, this is not so. Most of these various choices mentioned here hold together naturally, offering a robust, unified, and popular way of accounting for the character of perceptual experience. And there are heavy theoretical penalties, especially for the naturalistically minded, for adopting other options here, like Fregean or narrow representationalism, or a nonreductionist or restricted form of representationalism. So while our particular target here may seem overly specific, the worries we raise are quite general and apply to a major form of intentionalism, and can be generalized to other accounts of perceptual experience.¹¹

3 Representationalism about affect

First, a few preliminaries and some terminology. Our perceptual encounters with our environment typically result in the formation of perceptual beliefs. What makes these beliefs perceptual, in the way we want to highlight for the purposes of this paper, are two factors. One is their perceptual content: their content is the kind of content that perceptual experiences can have. The issues surrounding perceptual content are complicated. But in what follows we will focus on an uncontroversial kind of content: sensible qualities understood objectively. Surely, if experiences represent at all, they represent *at least* the sensible qualities that characterize the perceptual modality through which they are generated. The other factor stems from the fact that one can form perceptual beliefs in all sorts of ways. Here we are interested in the particular way

equally apply to the latter two. For a sustained criticism of representationalism about intransitive bodily sensations in general, see Aydede (2006, 2009, forthcoming).

¹¹ For instance, we think that the main part of our criticism below, when modified suitably, will also be effective against versions of naïve realism as well as certain weaker forms of intentionalism.

they are formed in direct causal response to perceptual experiences — their informational etiology. Our perceptual experiences typically non-inferentially prompt conceptual responses to the information they contain about their environmental objects and properties. These responses underwrite our ability to discriminate, identify, and recognize features of our environment. We can regard these direct conceptual responses (direct concept applications) as the non-inferential formation of perceptual beliefs — or, to highlight their occurrent character, perceptual *judgments*. These non-inferential beliefs or judgments are not only causally/informationally prompted by the experiences, but also, under appropriate conditions, epistemically justified by them. When it is important to differentiate these non-inferential tokenings of the perceptual beliefs/judgments with a particular sort of informational etiology, we will mark them as *direct* perceptual beliefs/judgments (cf. Dretske 1999). We will sometimes abbreviate these as ‘p-beliefs’ (or ‘p-judgments’), and concepts thereby applied, as ‘p-concepts’. It is important to keep in mind that perceptual beliefs are essentially first-order in the sense that they are about the extra-mental objects of experiences that directly prompt them, and not about the experiences themselves (they are not introspective). In other words, the relevant concepts involved in them are directly applied to what the perceptual experiences represent, and not to the experiences or to their phenomenal aspects.¹²

Before presenting the natural representationalist account of affective qualities, it is useful to start with a non-affective case. The contrast will make it easier both to understand representationalism’s main commitments and our criticism of it. Let *e* stand for a perceptual experience and *P* stand for whatever phenomenal aspect of *e* specifically represents a sensible property of an extra-mental object. Let *P(e)* refer to *e*’s being (or having) *P*.¹³ According to SR, when we see an object as red, for instance, our experience is representing the object as red.¹⁴

¹² To illustrate, suppose that I am in my study room and I hear my spouse in the kitchen telling my son not to touch the red apple on the counter. I form the belief: THAT APPLE IS RED. On our use this belief is not a direct perceptual belief — although it may be a perceptual belief if its content can be the content of a perceptual experience — because it’s not formed as a non-mediated direct response to a perceptual experience of an appropriate sort. Dretske sometimes calls this sort of indirect perceptual belief formation, displaced (epistemic) perception. Note that direct perceptual beliefs are psychologically mediated — by appropriate experiences, but not, intuitively, by other beliefs.

Although the terminology is somewhat ours, the clarifications and claims just made are common ground among intentionalists. In fact, Dretske (1981, 1995) and Tye (1995, 2000) take this capacity to form direct perceptual beliefs as a necessary condition for perceptual experiences to be phenomenally conscious. For instance, Tye’s requirement that the content of experiences be *poised* minimally requires this sort of capacity.

¹³ By ‘*P(e)*’ we don’t assume that experiences have intrinsic phenomenal qualities. Following our opponents, all we want to mark with this way of talking is that there is a way of characterizing experiences according to what it’s like to undergo them for their subjects. So, for instance, if representationalism is true, *P* is a representational property. Also we will be assuming that whatever further conditions presumed necessary for experiential representation (like being poised, non-conceptual, etc.) are in place.

¹⁴ We will use the property, *red*, as our *arbitrary* sensible quality that the objects of our experiences have. Similarly for the concepts, RED, REDDISH — see below. Later we will switch the example from red to bodily disorder as the latter is presumed to be a sensible bodily condition represented by the sensory-discriminative aspect of pain experiences.

The “reddish” phenomenology of our experience, P , is identical to (or entirely determined by) the property represented by our experience, in this case, *red*. Our experience of the object therefore has correctness conditions:

- $P(e)$ represents $\text{Red}(o)$, and is veridical IFF o is red.

So, when, on the basis of our experience $P(e)$, we form the *direct perceptual belief* that the object is red, then — bracketing certain skeptical scenarios — the truth value of our belief matches that of the perceptual experience. In general, if a perceptual experience is non-veridical, then so too will be the p-belief based on it, and vice versa. In other words, according to SR, the truth value of a p-belief generally matches that of the perceptual experience upon which it is appropriately based. Call this semantic parallelism:¹⁵

(PARALLELISM) p-belief(o is red) is true IFF $P(e)$ is veridical.

Finally, note that the *sensible qualities* of perceptual objects, *primarily* or *fundamentally*, qualify the objects of experiences — they are objective properties in this sense. If there is any sense in which they ever qualify their experiences, their so doing is only *derivative* (i.e., somehow depends on their qualifying the objects primarily).¹⁶ This is the analog of OV as applied to sensible qualities. Given SR+ST, we take it that this is all common ground. Now let’s state representationalism about affect.

3.1 Representationalism and the OBJECT VIEW (OV)

Representationalism with its commitment to the strong transparency thesis naturally invites the adoption of the OBJECT VIEW, according to which affective qualities such as pleasantness or

Also, following the more or less standard practice, we will use the capitalized words as names of concepts, where concepts are understood to be mental representations of a certain sort in more or less the psychologists’ sense.

¹⁵ Here we will ignore skeptical possibilities such as “veridical hallucinations”. These are not relevant to our present discussion. We will also ignore a particular line of representationalist response to inverted spectrum arguments, according to which Invert’s color experiences *systematically* misrepresent the actual colors of things she sees around herself despite the fact that her perceptual judgments about colors are correct. So, for instance, when Invert sees a ripe tomato in good light, her experience misrepresents its color as green — as having the same color as represented by Nonvert’s experiences of cucumbers — while her perceptual judgment, THIS IS RED, is true, when made as an appropriate response to her seeing the red tomato. This is a controversial response to a controversial scenario (even among representationalists), and is irrelevant to the way in which we will claim that PARALLELISM would be counterexampled if SR were to adopt the OBJECT VIEW — see below. So we set it aside here.

¹⁶ For example, we sometimes talk about sweet tastes, acrid smells, prickly or warm sensations. On SR, if there is any natural sense in which these qualities qualify the taste, smell, touch experiences, etc., this is only a *façon de parler* whose truth-conditions depend on whether these experiences correctly represent their object as being qualified by these qualities.

unpleasantness primarily qualify the objects of perceptual experiences and are represented by these experiences. The experiences get to have affective phenomenology in virtue of representing these affective qualities. Applied to pain, the affective quality, *painfulness*, primarily qualifies objective bodily conditions such as physical disorders or disturbances of various kinds occurring or about to occur in parts of the sensitive body that are also represented by pain experiences. So a pain experience typically represents *both* some kind of objective bodily disorder (sensory-discriminative aspect of pains) *and* its unpleasantness (pain's affective-motivational aspect). Let's use '*d*' to denote the bodily disorder and '*P*' to denote the sensory-discriminative phenomenology of pain experiences that represents the disorder.¹⁷ Also, let *P** be the negative affective phenomenology of a pain experience, *e*, that represents *d* as painful.¹⁸

The most fundamental tenet of SR is that *P** represents (just as *P* does) a non-mental objective property of the objects of experiences. This property, according to Tye (1995, 2000, 2005a, 2005b) and Cutter & Tye (2011), is the property of being bad to an individual to a particular degree:¹⁹

The content of a pain experience of an individual *A* is something like: there is a bodily disturbance of (physiological) type *d* in location *l*, and *d* is bad for *A* to degree *x*. (C&T 2011: 99)

C&T propose that the *naturalistic property* (*H*) to be identified with the property of being bad for *A* to degree *x* is

H = the property of being apt to harm individual *A* to degree *x*.²⁰

Thus, according to C&T, *P** represents *H* (=harmful, for short):

¹⁷ It is important to note that, for present purposes, we will be following representationalists in assuming that pain experiences are perceptual and represent (actual or potential) bodily disorders — this will be pains' *sensory-discriminatory* aspect. Here, our quarrel with strong representationalists is with their claims about pains' *affective-motivational* aspect, not with its sensory-discriminative aspect. Nevertheless, see Author (XXXX) for an argument that pain experiences are not perceptual and that SR cannot explain pain's sensory-discriminative phenomenology either.

¹⁸ Similarly, we will take *painfulness*, thus *P**, as an *arbitrary* affective quality for what follows.

¹⁹ Bain (2012, 2013) and O'Sullivan & Schroer (2012) also claim that the representational content of pain affect (*P**) is that *d* is bad. They are all strong representationalists but they leave open what kind of property *d*'s badness (*painfulness*) comes to — although they seem to have a naturalistic property in mind to be identified with badness.

²⁰ Similarly, C&T (2011: 105) identify pleasantness with the natural property of being apt to benefit individual *A* to degree *x* (= *H*). In what follows, '*H*' will refer to the *naturalistic non-mental* property (whatever it is) represented by the affective phenomenology, *P**, of experiences.

- $P^*(e)$ represents $H(d)$, and is veridical IFF d is harmful.

C&T comment on the notion of harm involved thus:

We can understand the notion of harm in relation to the notion of a teleological system. Very roughly, something harms a teleological system to the extent that it hinders that system (or one of its subsystems) from performing its function(s). (2011: 99–100)

What is important to keep in mind here is that, in both positive and negative affect, H is a *non-intentional physical property* or a *relation* that *doesn't involve e or anything mental* as a relatum, and is to be identified with the naturalistic property represented by the affective phenomenology of our experiences. So it is an objective property of the objects of experiences, not a property of the experiences.²¹

So far so good. According to OV, the affective quality, being painful, primarily qualifies the objects of experiences. This object is some kind of bodily disturbance, d , in the case of pains. C&T claim that the affective phenomenology of pain, P^* , represents d as being apt to harm its possessor to some specific degree (*harm* or *harmful* for short). So being painful is identified with being harmful, and is attributed primarily to d . But adopting OV in this way implies the perception-first view (PFV), which in turn implies that one cannot make the introspective judgment that

(I-JUDG) one's experience is painful,

unless one is capable of making the perceptual judgment that

(P-JUDG) d is painful.

²¹ It is H (or, “badness”) so understood that is the target of our criticism below. Although the particular value proposed by C&T under consideration seems perhaps the most natural candidate for H , we believe that our argument will generalize to any other proposed value of H so understood. Also, under OV, the representationalists could be seen as proposing that the affective qualities should be interpreted as a species of objective sensible qualities. However, because H is a highly relational (and possibly, partly but essentially historical) *second-order property of sensible qualities*, we are reluctant to describe representationalists this way.

So the attribution in P-JUDG is primary, just as OV says. Hence, given H , PARALLELISM implies that P-JUDG is false if d is not harmful — in other words, when $P^*(e)$ is not veridical.

But now consider the following situation. Let d be the removal of a scar tissue due to a serious burn by a doctor to prevent further pathological complications. So the removal (d) is not harmful. Then, $P^*(e)$ is not veridical.²² But then, given PARALLELISM, P-JUDG, when made by me, is not true.

But clearly, when sincerely made by me using the ordinary concept — what else? — PAINFUL, P-JUDG is true! We don't think we need to argue for this point at length, but one way to see this is to see that there is no (independent) evidence that would convince me that when I judge that:

(CONTRADICTION) Boy, this is good for me but so painful!

I am making a semantically *contradictory* judgment (where, 'this' refers to the removal, and good = beneficial/not-harmful). If the property attributed in P-JUDG were identical to the property of being harmful ($=H$), CONTRADICTION's truth-conditions couldn't obtain in any possible world — formally spelled out: 'this is both harmful and not-harmful'). But clearly I would insist on its being true and balk at the suggestion that I am in fact making a contradictory judgment. And this is the ordinary default situation (not just me). Note that these kinds of judgments aren't rare in the case of first-order affective judgments — there are *plenty* of them we encounter in everyday life. Take, for instance, just one tiny portion of the plethora of daily examples: parents' constant pleadings like "eat/do x , it is good for you", where x is also judged to be unpleasant). If P-JUDG and its ilk attribute H to the objects of our experiences, then a lot of the situations we are commanding, pleading, praising, recommending, judging, etc., are literally contradictions.

Can representationalists say that the identity between, say, painfulness and harmfulness is *a posteriori*? To see that this doesn't help in this context, consider other contradictory/false beliefs arising out of failure to believe an *a posteriori* identity: when the reference of the terms/predicates in these identities are independently traced or established, we readily withdraw our judgments on ground of realizing that they are false or contradictory. For instance, when Pierre learns that London=Londres, he would no longer assent to "Londres is pretty but London isn't" (Kripke 1979). Same with contradictory judgments using indexicals, demonstratives, or names. Same with judgments using natural kind terms. But what evidence could convince me (or, any person who is not a dogmatist about SR) that, when I judge CONTRADICTION, I am in fact expressing a contradictory belief? The challenge here for the representationalists who might be

²² This (and, in general, the frequent non-veridicality of all kinds of affective experiences) is readily admitted by Cutter and Tye (2011), and other representationalists — see below.

tempted to claim *aposteriority*, then, is to come up with *independent* evidence for the truth of SR+OV about affective qualities. We don't see how they can do that. But, in fact, they don't even try, since, as far as we can tell, they already grant that affective qualities are not to be identified with *H* (see below).

Furthermore, insisting on the *a posteriori* identity and therefore the actual falsity of P-JUDG would amount to changing the topic. A representationalist account that would insist on systematically correcting people's judgments about what/when they find things painful/pleasant would not be an account of experiential affect as this notion was originally targeted for explanation. Recall that the representationalist project is to reduce all experiential phenomenology to representation (of a certain sort). SR agrees that many experiences have an affective phenomenology. The professed project, therefore, is to understand the pleasantness or unpleasantness of experiences in representational terms. The explanandum here *is* a phenomenon conceived in a certain way. This way gives the experiencers a certain sort of cognitive and epistemic authority regulated by the phenomenology of their affective experiences. Even if we grant that the affective phenomenology (mis)represents *H*, our conceptual responses to this phenomenology obviously don't follow its purported representational norms and function in the way SR envisions. An account like SR+OV that implies that they do is empirically false. An account that suggests that they ought to is at best unmotivated, or at worst radically revisionist in a way that misses its original explanatory target. Anyway, we don't think we need to elaborate this any further: as said, almost all representationalists reject the identity of affective qualities with *H*.

3.2 Representationalism and the EXPERIENCE VIEW (EV)

So contrary to natural expectations, OV isn't the right view for representationalists. And, indeed, Tye (2005a,b) and Cutter and Tye (2011), along with almost every other representationalist, reject OV about affective qualities. Instead they hold what we have called the EXPERIENCE VIEW (EV), according to which affective qualities *primarily/fundamentally* qualify perceptual experiences rather than their objects.

One way to motivate EV is to observe that given *H*, the properties attributed in I-JUDG and P-JUDG can't exactly be one and the same property: experiences aren't harmful in the same way their objects can be harmful. We have made the same observation before in the context of pleasantness when we were setting up the framework for the locus of affect attribution. Similarly, it is not implausible to say that 'painful' in I-JUDG and P-JUDG is ambiguous — irrespective of the nature of *H*. Indeed, following common sense, Tye finds it quite plausible that:

... when we say that something is painful, what we normally mean is that it *causes* the feeling of pain. Thus, we speak of a cut, a bruise, a sore, an operation, a cough, a lashing as painful, that is, as causing pain. This is why, if a person feels a pain in his left arm,

and his doctor informs him that the cause of the pain lies in a disturbance in his heart, we allow that what is *really* painful here is the disturbance in my heart. If the doctor has made no mistake, *that* is what is actually hurting the patient. It is evident that [pain experiences] themselves are not painful in this [derivative] sense. Pain does not cause the feeling of pain. (Tye 2005a: 114)

So, given the proper disambiguation, in P-JUDG, *d* is painful in the sense that *d* is *causing a painful experience*, and in I-JUDG, the sense of ‘painful’ is primary. So contrary to first-impressions, judgments like P-JUDG that attribute affective qualities to the *objects* of affective experiences don’t track a completely objective property of these objects such as their being harmful or beneficial. In this regard, the attribution of affective qualities in the act of experiencing them is radically different from the parallel attribution of sensible qualities to the objects of our experiences.

As should be expected, this view comes with an INTROSPECTION-FIRST VIEW (IFV), which implies that affective qualities that primarily qualify our experiences (=EV) are introspectively knowable directly, i.e., *not* by an epistemic detour through our capacity to make direct perceptual judgments about the objective qualities of their objects. In other words, when we make introspective judgments like I-JUDG, our attribution of affective qualities to our experiences does not depend on our capacity to have perceptual judgments about the relevant range of objective properties of their objects.

However, this seems to be a clear violation of the *strong transparency* thesis (ST): it implies that we can come to know a quality *of* our experiences without necessarily coming to know a corresponding quality of the object of our experiences represented by the former quality. But if ST is genuinely counter-exemplified, SR is refuted. Given the untenability of OV, EV is the only option for representationalists. And in fact, they routinely seem to take that view. But EV violates ST. And if ST falls, SR falls. As far as we can tell, representationalists, including Cutter and Tye, haven’t yet fully appreciated the true impact of adopting EV about affective qualities. It’s quite puzzling why. Here is a guess. Like many others, C&T correctly think that even if my experience misrepresents *d*, it is still painful. So, on this basis, it would be natural for me to come to believe that

(FEELS-BAD) *d feels bad/painful.*

This is not, of course, of any immediate help to explain the difficulty with adopting EV since this is not the content of a first-order p-belief anymore. It is a belief about how *d* is *experienced* by me. It’s an introspective belief similar to I-JUDG in the sense that it is essentially a commentary

on my *experience* — albeit as an experience of *d*. FEELS-BAD is ambiguous between two readings:

(FEELS-BAD_E) I am painfully feeling *d*

and

(FEELS-BAD_O) I am feeling *d* as painful.

These readings would roughly correspond to adopting EV and OV about painfulness respectively. When C&T make claims like FEELS-BAD, they may have FEELS-BAD_E in mind but somehow think that it is compatible with SR on account of the availability of the second, FEELS-BAD_O, reading. The trouble, as we have seen, is that neither reading is of any help to explain how they can hold EV and think that ST is preserved — while they seem to reject OV.

Thus neither OV nor EV is an option for representationalists: they can't answer the locus question about affective qualities. But as far as we can tell, there isn't any other option for representationalists. Nevertheless, the next section presents a likely response that they might try.²³

4 Reformulating representationalism about affective qualities

Here is how it would go: painfulness is indeed a phenomenal feature of pain experiences. This feature is not to be identified with harmfulness. Indeed experiences are not themselves (let's grant) harmful. It's what they represent, *d*, that is harmful, and thus, *d* is only derivatively

²³ The next section is inspired by passages like this:

“...phenomenal properties of experiences, i.e., properties constitutive of what it's like to undergo those experiences, are identical to certain representational properties of those experiences, in particular properties which are constitutive of *what* the experience represents (as opposed to, say, intrinsic properties of the vehicles of representation), such as the property of ascribing redness to something, the property of representing loudness, the property of having the content that *p*, etc. And most versions of representationalism at least accept the weaker claim that for every phenomenal property *P*, there is a representational property *R* such that, necessarily, an experience has *P* iff it has *R*. Take, for example, the phenomenal property — call it phenomenal reddishness — distinctive of experiences as of red things. Plausibly, the representational property corresponding to reddishness, i.e., the representational property *R* such that necessarily, an experience has phenomenal reddishness iff it has *R*, is the property of representing something as red. But now consider the property of pain experiences — call it painfulness — that we allude to when we say that a pain in the leg feels bad. What is the corresponding representational property in virtue of which our pain experiences are painful?” (C&T 2011: 92–93, italics ours).

Note the infelicity in the opening line: the properties constitutive of *what* an experience represents cannot be the properties such as “the property of ascribing redness to something, the property of representing loudness, the property of having the content that *p*, etc.” Such infelicities abound in the representationalist literature. Nevertheless, what follows is a way of developing the idea that a pain experience has the intentional property of *representing-harm* and it is *this* quality of pain experiences that we directly introspect when we introspect the painfulness of pain experiences.

painful. So EV about affective qualities is correct. In our evolutionary history, we have, for some reason, picked up a way of thinking/talking about affective experiences so that we attribute the affective qualities primarily to the experiences, rather than to their objects. But one can still argue that the property that we attribute to experiences when we judge them to be, say, painful is their *representing* something as harmful to the experiencer. So there is the representational vehicle, the pain experience (*e*); there is also what it represents, bodily disturbance (*d*) and its being harmful. Then there is the intentional relation between the two: the vehicle *e*'s representing *d* as harmful. Because it is an intentional relation, *e* can represent something as harmful even in the absence of that "something". So *e*'s representing *d* as harmful is a property of *e*, and it's properly intentional as required by SR. To explain how *e* could come to have this property is the goal of various naturalization projects in psychosemantics. But for now, the proposal is that painfulness (or affective qualities in general) is to be identified with the representational property of *e*: namely, *e*'s *representing* something as harmful. So:

(PH) *e*'s painfulness = *e*'s representing-something-as-harmful-to-the-experiencer-
A-to-degree-*x* (*e*'s *representing-harm*, for short).

This claim should not be confused with:

d's painfulness = *d*'s being apt to harm individual *A* to degree *x*
(*d*'s *being harmful*, for short).

SR does not propose to identify affective qualities with objective (at least, non-mental/non-intentional) properties of objects of experiences (i.e., with *H*). So, following EV, it might be claimed that affective qualities are special in some way. This is why we have separate concepts and words for them that are attributed/applied to the affective experiences themselves, rather than to their objects. But what is being attributed (an intentional property as per (PH)), a representationalist might claim, is still perfectly in line with SR and naturalistically kosher — assuming we have a successful naturalistic psychosemantics.

For comparison, consider the visual experiences that represent red things. There is a particular visual phenomenology to seeing red. Outside philosophy classes, usually we don't have ordinary concepts/words to refer to this phenomenology. We say things like: the way red things appear or look to us. But, if we wish, we might introduce quasi-technical terms, like 'red*', 'red-prime', or 'reddish' experiences to refer to the particular phenomenal feature "distinctive of experiences as of red things" (C&T 2011: 93). By stipulation, then:

(STIPULATION) The property of being reddish = the phenomenal property distinctive of experiences as of red things.

The representationalist proposal then is this:

(RR) e 's reddishness = e 's representing-something-as-red (say, as having a particular surface spectral reflectance profile).

When we have a naturalistic psychosemantics to explain the representation relation in question, we will solve the problem of experiential phenomenology, and hopefully, the problem of phenomenal consciousness. So what is the problem?

One of the immediate problems with this response is that it seems to violate ST. Recall:

(ST) Any quality that we (can) epistemically encounter in the introspection of an experience is a quality only (widely) *represented* by this experience, thus not a quality *of* the experience (*a fortiori*, not an *intrinsic* quality of the experience).

If we accept the suggestion that the property of representing-harm is a property *of* the experience e and thus not widely represented by e , then, to preserve ST, we need to conclude that we don't (can't) epistemically encounter this property in introspecting e . But this is absurd: under the proposed new formulation of SR, the point is precisely that this is a property *of* e and we can and do directly become aware of this property. The spirit of the proposal is that this property is still an intentional property, and so, is not an intrinsic property of e — hence the thought is that it's kosher for representationalists to advert to such properties in introspection. So, at a minimum, we need to reformulate ST following the above reformulation of SR. But it is not at all clear how to do this in a way that will still support SR over weaker forms of intentionalism. Here is one way it may go consistent with the reformulation of SR (based on PH and RR above) — we may call it WEAK TRANSPARENCY:

(WT) Any quality that we (can) epistemically encounter in the introspection of an experience is an intentional quality of this experience,

where, for instance, reddishness and painfulness (as fixed by PH and RR above) are among the introspectible intentional qualities of an experience. But this is *too* weak. It is compatible, for instance, with forms of intentionalism that are not (strongly) representationalist: intentionalist

theories about phenomenal character that are Fregean or narrow or impure would endorse WT, as would the (non-reductionist) qualia theories that take a phenomenology-first approach to intentionality. This is not accidental. If ST follows from SR, then, of course, any transparency thesis weaker than ST (that may be compatible with EV) will also follow from SR. Thus WT may be true but irrelevant. Remember that we have reformulated SR in the above way *because* we were trying to understand how adopting the EXPERIENCE VIEW about affective qualities is compatible with ST. The representationalist suggestion on the table now is that the affective qualities we attribute *to* experiences (*as per* PH) *in fact* intentional properties. But this suggestion, or WT for that matter, doesn't help us to solve the original problem, namely, how ST and EV can be compatible — unless it is a tacit admission that they are not.

The main problem with the suggested reformulation of representationalism is that there is a very crucial difference between (PH) and (RR): grasping the property of being reddish in introspection is “transparent” in a way grasping the property of being painful is not. If so, we still don't have an explanation of how EV and ST can be compatible. Let us explain.

5 Introspective knowledge of affective qualities

To start, note that the identity proposed in (PH) between *e*'s painfulness and *e*'s representing-harm is meant to be a *theoretical identity* and it is by no means obvious; in fact, in some ways, it is quite perplexing. For, to begin with, painfulness doesn't seem to be an intentional property at all: if (PH) is true, then when we judge our experiences as painful (pleasant, etc.), we are, *as it turns out*, attributing an intentional property to them that makes them answerable to semantic evaluation. So in the above example, when I judge my pain experience to be extremely painful when, on SR, it misrepresents the removal of scar tissue on my face as harmful, I am in fact attributing an intentional property to my experience that makes it, on this occasion, non-veridical. Despite my (conceptual level) knowledge that the removal is not harmful, and in fact, beneficial, I am not (unless I am an avid representationalist) in the least bit inclined to attribute falsehood to my painful experience; more dramatically put, I am not in the least bit inclined to blame my experience for deceiving me or telling me a lie. In other words, despite my finding my experience very painful, I don't at all find it as (mis)representing the removal of scar tissue on my face as harmful. We take this observation to be true for most ordinary people who have the relevant concepts. This observation doesn't immediately refute SR, of course. For such epistemological or psychological incongruities are not uncommon in theoretical identities that may sometimes involve surprisingly incongruent concepts flanking the identity sign. But we think that *this* “theoretical” identity is special and requires additional independent evidence given the general motivation behind defending SR and SR's general commitments.

Consider (RR). ‘Reddish’ is a technical term (just as ‘red*’ or ‘red-prime’ are), not ordinarily used to attribute a phenomenal property to one's visual experiences. Nevertheless, it is, let's grant, a well-defined term stipulatively introducing a more or less well-behaving property (well, let's assume it does). What is the status of (RR)? Defenders of SR intend to take (RR) and the like also as an *a posteriori* theoretical identity. But clearly there is a difference

between (RR) and (PH). If (PH) and (RR) are true, we are introspecting the property of representing-harm and the property of representing-red when we introspect our relevant experiences. But, on SR, we need perceptual concepts to make introspective judgments about these experiences.

Given how we have introduced the concept REDDISH, we can come to know that our experiences are reddish only if we have the concept RED that applies to red things that these experiences represent. Given the general motivation behind SR, this is as it should be: our perceptual experiences are strongly transparent in that anything phenomenal we can come to know about them is, somehow, via knowing *what* they represent. Indeed, according to SR, when I look at a red bulgy tomato in good light, I know that it (phenomenally) looks red to me — indeed, I know what it is like to experience the redness of the tomato — only if I can apply a first-order perceptual concept, RED, to *what* my experience represents. The p-concepts like RED involved in this sort of introspective knowledge may be special in that their acquisition may require serious epistemic/psychological constraints on how they depend on the experiences from which they are directly and immediately acquired. Whatever the case may be about how these p-concepts are acquired and deployed, according to SR, they apply not to the experiences but to what these experiences represent. This is all common ground.²⁴ Given all this, there appears to be no special puzzle about how I acquire the *introspective knowledge* that my experience is reddish or that something looks red to me.²⁵

²⁴ For the clearest and very emphatic statement of SR's commitment to the availability of p-concepts for introspection, see Dretske (1995: 138–140) and (1999: 18–20).

²⁵ Well, along with representationalists, let's assume there isn't — we are granting this only for the purposes of the present dialectic. We, in fact, think that there are serious problems with any view of introspection compatible with SR+ST.

Here we don't discuss representationalists' preferred theory of introspection of experiences that sometimes goes by the name of 'displaced perception model' (DPM) baptized and initially developed by Dretske (1995, 1999). DPM is an *inferentialist* theory of introspection, according to which introspective knowledge is not direct at all but inferred from a certain sort of first-order perceptual knowledge with additional premises. See Aydede (2002) for a criticism of this view. Tye has declared his allegiance to this model in many of his writings (e.g., Tye 2000), but he rejects inferentialism. Frankly, we have difficulty understanding Tye's account of introspection. In fact, as far as we can tell, all representationalists seem to be having difficulties with deciding on a suitable theory of introspection, or else keep silent on the topic. Byrne (2009, 2012), for instance, further develops the inferentialist view. But as far as we can see, he appears to save Dretske's version at the expense of the existence of experiences. It appears that it's beliefs all the way down for Byrne. So he ends up denying that there are experiences in the sense that makes them philosophically interesting as above and beyond the puzzles that attach to beliefs. Besides, his positive account doesn't work — see Aydede (2002). Dretske has given up and become a skeptic about self-knowledge and he confesses that it is a "mystery" how we acquire introspective knowledge at all (see, for instance, his 2003, 2006, 2012). Tye's account in his recent work (2009) has become increasingly more difficult to follow. We think none of this is accidental.

Luckily, though, we don't need to discuss DPM or its variants. All representationalists under discussion in this paper, recall, are phenomenal externalists and believe in some form of transparency; accordingly, they think that the introspection of an experience of a sensible quality *F* requires the availability of the p-concept of *F*, and that's all we need in this paper. They differ in their commitments to how this concept is deployed in the process of introspection. DPM says that this concept is not *directly* applied to any quality of the introspected experience, but is used rather to describe what property is being represented in the experience. (Note that to say that p-concepts are required for the introspection of experiences is not to say that we can't experience/perceive the represented sensible

Given (STIPULATION), when I look at the red tomato in good light, it is *a priori* that

- I know that my experience is *reddish* only if I know that my experience is of the sort distinctive of experiences as of *red* things.

So there is literally a conceptual connection between the concepts REDDISH and PHENOMENAL-PROPERTY-DISTINCTIVE-OF-EXPERIENCES-AS-OF-RED-THINGS, and, for that matter, REPRESENTS-RED, in that they can't be possessed without possessing the concept RED. So,

- Necessarily, one has the concept REDDISH only if one has the concept RED.

We take it that what we have said so far about the introspective knowledge of our experiences of *sensible* qualities is all common ground between us and representationalists.

But, on the representationalist proposal we are now considering, the introspection of *affective* qualities works radically differently. It is not *a priori* that

- I know that my experience is *painful* only if I know that my experience is one of those distinctive as of *harmful* things.

If I don't have the concept HARMFUL or I have additional empirical information of a certain sort, it may not be even true. As should be obvious, there is no conceptual connection between the concepts PAINFUL and HARMFUL, and for that matter, REPRESENTS-HARM. So it is just not true that:

qualities without these p-concepts: on this view, without p-concepts we are blind only to our experiences, not to the world).

Interestingly, Tye, until his 2009 book, had also been defending so-called phenomenal concepts. We think that the way phenomenal concepts are currently understood among physicalists is incompatible with the fundamental tenets of representationalism — what follows in this section lays out the gist of how an argument for this claim would go. It is tempting to speculate that this might be part of the reason why Tye in fact gave up phenomenal concepts in his (2009). It's interesting to observe that no other representationalists have ever been friendly towards phenomenal concepts. (Lycan might seem to be an exception but he is not a pure and non-restricted intentionalist, thus not a strong representationalist in our sense: for example, Lycan defends a *functionalist* account of affective qualities and of phenomenological differences among modalities — see his 1987, 1996. Furthermore, his introspection-first view about what makes perceptual states phenomenally conscious states — his HOP account of conscious states — is seriously at odds with SR.)

- one has the concept PAINFUL only if one has the concept HARM (or, HARMFUL, or REPRESENTS-HARM).

These concepts are distinct and conceptually unconnected. One is a monadic concept and the other is complex. Their acquisition and application conditions are quite different and independent. But to say that they nevertheless attribute one and the same property doesn't help to remove the mystery of how EV and ST are compatible.

If SR+ST is true, it is *not* accidental that the introspection of experiences requires the availability of the very same p-concepts that apply to the qualities of the objects of experiences: REDDISH is, after all, a technical term conceptually manufactured out of the concept RED (see above). So it is not accidental that there is an *a priori* connection between RED and REDDISH (i.e., EXPERIENCE OF RED or REPRESENTING-RED): it is required by SR+ST. On this view, the only sense in which I know that my experience is reddish is the sense in which I know that my experience is an experience *as of* red, or that it represents red. If I don't have the p-concept that applies to the object of my perception (not to say anything about concepts such as EXPERIENCE, REPRESENTS, etc.), then I can't have the concept required to introspectively believe *what* experience I'm having. In brief, one can't track an intentional property *qua* intentional without representing to oneself what the intentional property represents — indeed without one having any clue whether the property in question is even intentional. But, if so, insisting on (PH) would amount to admitting that there are qualities of experiences that we can *directly* become aware of (i.e., we can “i-conceptually” track) without, strangely enough, becoming aware of (i.e., without “p-conceptually tracking”) the qualities represented by these experiences. So adopting EV and thus the INTROSPECTION-FIRST VIEW (IFV) for affective qualities threatens the core of representationalism by making strong transparency optional.²⁶

Here is another way to make our point. Recall that when I make the introspective judgment:

(I-JUDG) my experience is painful,

I am *not* attributing the property of being harmful to my experience in exercising my concept PAINFUL — as agreed by all parties, painful experiences aren't harmful in the intended sense. Furthermore, whatever property I am attributing to my experience, that property is not conceived by me in a way that requires me to have a p-concept that applies to what my painful experiences (by being painful) represent. I-JUDG is not an inferential or abductive judgment: it is introspective and seems to be the immediate result of *directly applying* my i-concept, PAINFUL,

²⁶ In a nutshell, this is why the defenders of SR+ST insist of the availability of p-concepts for introspection (roughly, PFV). Cf. Tye (2005a: 116) for a response to Aydede (2001) that seems to miss this point.

to an instance of an affective quality of my experience. This means that I am introspectively capable of responding to a quality of my experience *without having the slightest clue what that quality may represent, or even whether it represents anything*.²⁷ It is a mystery on SR how we can *directly* apply i-concepts like PAINFUL to our experiences without having the slightest idea that this quality may be representational (let alone *what* it might represent). This is just the denial of strong transparency that representationalists are committed to.

Indeed, it is very natural to use the following form, instead of I-JUDG, in expressing my introspective judgment:

(I-JUDG*) My (pain) experience has *this* quality,

where the intended demonstration is clearly to an instance of painfulness instantiated in my experience. Here the demonstration can't possibly be to an instance of harmfulness: as said, my experience isn't harmful in the way its object (*d*) might have been harmful. I may go ahead and add to I-JUDG*, "it's painful" or "it's unpleasant" — expressing a fact about how I conceive of this quality. Clearly, this seems to be an act of inner demonstration involving at the same time an act of direct recognition and thus categorization, which suggests that I am *directly* applying a "recognitional" concept, PAINFUL or UNPLEASANT, to my experience. We have already seen that I cannot be applying this concept to *d* either — remember we are trying to understand how ST can be compatible with EV.

So the suggestion that EV is true of affective qualities but that affective qualities are *a posteriori* identical to intentional qualities of experiences in the way proposed by (PH) turns out to be *not* compatible with the strong transparency (ST) claim that representationalists are committed to. But if ST is falsified by affective qualities, so is representationalism. Given that the representationalism under review is non-restricted (i.e., is meant to apply to *any* kind of phenomenal character), SR is just false.

But we are still not finished. Cutter and Tye package their version of representationalism with the tracking theory of intentionality (TTI), as they call it, which aims to explain the naturalistic underpinnings of how the pain experiences come to represent *H* (= being harmful or beneficial). We will see that this is misleading. Their version actually is a brand of teleological psychosemantics rather than an informational (tracking) one, and this makes a difference in evaluating alternative theories. We now turn to its critical discussion.

²⁷ Compare the following passage: "On my account, what it is exactly that a given experience or feeling represents need not be accessible to the subject's cognitive centers, including his or her powers of introspection, except in the most general and uninformative way (for example, as an experience of *this* sort). Nonetheless no two states that differ in the relevant representational contents can differ phenomenally, I claim; moreover, any two states that are alike in the relevant contents must be alike phenomenally. Why? Because phenomenal character is one and the same as representational content of the appropriate sort" (Tye 1996: 52). It's not clear to us at all what kind of concept may be expressed by '*this* sort' in 'experience of *this* sort' consistent with the demands of ST.

6 Tracking

Let's start with some observations about the nature of the particular proposal C&T are making regarding *H*:

H = the property of being apt to harm individual *A* to degree *x*.

First, this is a *second-order* property in the sense that tokens have this property in virtue of having another (first-order) property (or set of properties). For convenience, we can identify this first-order property with the property of being a bodily disturbance (*BD*). Again, given that the character of the proposal is meant to be generalizable, other proposed values for *H*, e.g., being beneficial to an individual *A* to degree *x*, would similarly be second-order: tokens would have them in virtue of having first-order sensible properties objectively understood.

Second, *H* is essentially relational in at least two senses. First, it is essentially a relation between the instantiation of a first-order property *BD* and the individual whose body instantiates it. Its metaphysical nature is comprised of a certain interaction between two relata: the instantiation of *BD* and how it effects a certain individual. Second, it is essentially a *historical* property. Necessarily, the first-order property *BD* is harmful to individual *A* only if *A* is a teleological system. Necessarily, *A* is a teleological system only if *A* has a particular sort of evolutionary history (or designed by such — we will omit this).

Clearly, even if *H* is a natural property, it is a highly complex property. Is it a *trackable* property? That depends on what it takes to track a property. Tracking theories of representation have been traditionally associated with information-theoretic or indicator semantics and also traditionally distinguished from teleological or consumer semantics.²⁸ We will now show that C&T's preferred semantic theory about how experiences represent *H* (i.e., have affective phenomenology) is not a tracking but a teleological one — contrary to their own advertisement. We will then argue that going teleological on this score is not superior to alternative theories (such as old-fashioned psychofunctionalism).

C&T characterize their version of “tracking theory of intentionality” (TTI) in the following way:

According to TTI, tokens of a state type *S* in an individual *x* represent that *p* if and only if: (7) Under optimal conditions, *x* tokens *S* iff *p* and because *p*. (2011: 100)

²⁸ For useful reviews, see Adams and Aizawa (2010) and Neander (2012).

Optimal conditions are those “conditions we were designed to operate in by natural selection” (102). Given this, C&T give the following formulation for what it is for an individual *A* to undergo a pain experience (we paraphrase and revise — cf. p. 100):

According to TTI, an individual *A* tokens a state *S* (= a painful pain experience that strongly represents that there is a disturbance *d* & *d* is apt to harm *A* to degree *x*) such that

(8) under optimal conditions, *A* tokens *S* IFF

- (a) *S* is a functional state type *F* defined (very roughly) by its forward-looking causal role of bringing about *d*-avoidance behavior,
- (b) there is a disturbance *d* and *d* is apt to harm *A* to degree *x*,
- (c) *A* tokens *S* because: there is a disturbance *d* and *d* is apt to harm *A* to degree *x*.

C&T interpret “because” as a causal dependence relation.²⁹ The first conjunct of (b) is supposed to capture the representational content of the sensory-discriminative aspect of a pain experience, and the second conjunct the representational content of the affective-motivational aspect.

In their original formulation, C&T give only (a) and (b), but then say “[w]e suggest that the value of *S* satisfying (8) is a functional state type *F* defined (very roughly) by its forward-looking causal role of bringing about ... avoidance behavior.” (2011: 100). Clearly, this condition needs to be added to the “tracking” formulation — as above. In fact, not only that, the avoidance behavior needs to be specific to *d* — otherwise, it is easy to find counterexamples to (8) above.

Now take *d* to be the removal of scar tissue on *A*’s face now, and *e* to be *A*’s pain experience now being caused by this. Further, assume, as above, that *d* is not harmful (in fact, beneficial). Is *e* a token of *S*? Hard to tell. Clearly (b) is not satisfied. Also clearly, (c) is not satisfied since the painfulness of *e* (*P**) is not now being caused by *d*’s being harmful simply because *d* is not harmful in this situation. A non-existent condition cannot cause any event. Given that *e* is a token of a painful pain experience, C&T have no option but to say that the conditions are not optimal. Indeed, in discussing pain causing disturbances that are not harmful (medically necessary needle punctures in skin or leg amputation), that is what they say:

In these cases, the resultant experience represents the bodily damage in question as bad for one, when in fact it is not bad for one in the circumstances. These cases are not a

²⁹ See their discussion on p. 101. However, in the next page, they seem to distance themselves from a causal dependence interpretation of ‘because’. We don’t know what to make of this. We will follow their initial claim that “[i]n the case at hand, however, we don’t think that there is any plausible alternative dependence relation that could account for the truth of the relevant ‘because’ claim” (2011: 101).

problem for our view, since under optimal conditions (roughly, the conditions we were designed to operate in by natural selection), getting one's leg cut off or having one's skin punctured is bad for one. It is because of this fact that, when we undergo such tissue damage, we token functional states characterized by their forward-looking functional role of bringing about aversive behavior. And this, on our theory, is what makes it the case that these token states represent the disturbance in question as bad. None of this, however, requires that the disturbance is bad for one in one's actual present conditions. (2011:102)

So the painfulness of *e* misrepresents *d* as harmful. The reason is that the conditions aren't optimal, which is to say that the teleological system *A* is not operating under the conditions in which *A* is designed to operate by natural selection. To say this, on C&T's view, is to say, among other things, that *e* belongs to a type of state *S* whose tokens were caused by disturbances that were harmful for the ancestors of *A*. This in turn means that in *A*'s ancestors bodily disturbances like *d* caused tokens of a state type which produced *d*-avoidance behavior which then enabled *A*'s ancestors to survive and reproduce with the trait to produce tokens of state type *S* with functional role *F* specified in (a) above. Survival was the causal result of avoiding *d*-like disturbance. Note that this is the only sense in which the harmfulness of disturbances like *d* were causally relevant: in our evolutionary history, there were ancestors who actually produced *d*-avoidance behavior in causal response to *d*-like disturbances which enabled them to survive, reproduce, and pass on the trait of producing similar behavior in similar circumstances. The 'because' in (c) of (8) gives the impression that it is the 'because' of nomological causal covariance as required by tracking theories — namely a nomological dependency between the instances of being harmful and the tokening of *S*. But it turns out that it is the 'because' of historical natural selection which is meant to express the causal explanation of why the type *S* with functional role *F* was selected in the past.³⁰ So the causal dependency of *S* tokens on the harmfulness of *d* consists in nothing but the historical fact that if our ancestors had not produced *S*-generated avoidance behavior in response to *d*-like states, they would have perished. There is nothing else to the 'because' of (c). We are not saying that harmfulness thus understood cannot be a causally relevant property. Despite some controversy surrounding the issues, we are prepared to grant that it can (it is). But it is misleading to present the causal efficacy of harmfulness of *d*-like disturbances understood this way as the causal efficacy (nomological dependency) of properties that figure in tracking theories. To see this, compare pain affect with what the sensory-discriminative aspect of pain experiences represent (call this phenomenal aspect *P*).

³⁰ Even this claim is controversial. There is a lively and ongoing debate in philosophy of biology whether traits can be causes. But we will ignore this. Compare our point to Dretske's distinction between structuring versus triggering causes (Dretske 1988).

Bodily disturbances that give rise to pain experiences with *P* are those that are detected by proper sensory mechanisms. In the case of pain, these are tactile and nociceptive sensory modalities. These detect certain physical features of the relevant range of stimuli, such as the kind of tissue disturbance (thermal, incisive, pressure, etc.), their location in the body, their volume, intensity, duration, etc. Call these special physical features of the nociceptive stimuli, *N*. So a rough approximation of a tracking theory of sensory representation would say something like:

Individual *A* tokens a state type *S* in *A* (= a token pain experience with P-phenomenology which represents that *N* is occurring in *A*'s body) IFF under optimal conditions *C*,

- (a) *N* is occurring in *A*'s body and
- (b) *A* tokens *S* because *N* is occurring in *A*'s body.

The 'because' here can have and has standardly had the sense supported by nomological causal co-variance. There is then a variety of ways in which misrepresentation can be accounted for in terms of the non-optimality of conditions. One standard way of going about this is to say that *S* is a state type of sensory modality *M* whose function it is to *indicate* *N* (in the information-theoretic sense) and that the present conditions hinder the proper functioning of this modality (Dretske 1995). Note that even though the biological function may need to be explained historically by natural selection, what fixes the content is what is supposed to be indicated by *S* tokens, namely instantiations of *N*. This requires that *N* is the kind of natural property that can be detected by a sensory modality.

Now consider the *painfulness* of pain experiences, $P^*(e)$, which is supposed to represent harmfulness of *N*. What is the naturalistic psychosemantic story? Harmfulness, *H*, as mentioned before, is a second-order property of *N*. Most instances of *N* are harmful, but not all, in fact some are beneficial. C&T's story, as we have seen, tells us roughly that tokens of *S* (= $P^*(e)$) represent harmfulness of *N*-tokens on the ground that *S* was selected because of its role in producing *N*-avoidance behavior in the past. Presumably, those historical tokens of *S* that resulted in the survival of their possessors *indicated* (carried information about) the occurrence of *N*-tokens. Did they also *indicate* that these *N*-tokens were harmful? No. There was no natural property apart of *N* that nomically controlled the tokenings of *S* in those historical contexts. This is not to say that *N* wasn't harmful then.³¹ But it is to say that their harmfulness was not the causally relevant property that explained the tokenings of *S* *then*. It was the downstream causal effects of *S*-tokenings, which resulted in the survival and reproduction of the

³¹ What made *N* harmful in these historical contexts? Simply the fact that those who suffered *N* and didn't produce *N*-avoidance behavior tended to perish. This is a second-order relational property of *N* partly constituted by *N*'s causal consequences upon those who didn't behave in a certain way. Clearly, not a property whose detection nomically controlled *S*-tokenings, thus not a property that *S*-tokenings carried information about.

possessors of *S*-tokens, etc., that explain why *S*-tokenings were selected and passed on to future generations. It is to this fact that C&T advert to in their naturalistic account of how *S*-tokens represent harmfulness of *N*-tokens *now*. But this is not a tracking psychosemantics, it is a form of consumer-oriented (functionalist) teleosemantics. What fixes the content of painfulness of pain states is the downstream forward-looking functional role that these states play and how this role was selected in our evolutionary history. These roles are essentially causal/functional roles.

So it turns out that (8) is an unhappy way of offering two different kinds of naturalistic psychosemantics for the representational content of pain experiences. The sensory-discriminative phenomenology (*P*) represents bodily disturbance, *N*. The affective-motivational phenomenology (painfulness of pain, *P**) represents *N*'s harmfulness. The former is explained by an informational (indicator) psychosemantics (a backward-looking nomic causal account), and the latter by a teleological psychosemantics (a forward-looking causal account).³²

This is not an objection to C&T's representationalism about pain phenomenology. But it is an objection to their misleading claim that their representationalism about the painfulness of pain is a tracking one. The painfulness of pain represents harmfulness not because it indicates harmfulness, but because pain experiences have historically played and to some extent still play a certain functional role — essentially involving avoidance behavior.

7 Representationalism about pain affect is unmotivated

We will now argue that this sort of SR about the pain affect is unmotivated at best because *that* the painfulness of pain *represents* harmfulness (*H*) doesn't do any explanatory work and thus is not needed beyond saving the letter of SR. The gist of our argument is simple. If the painfulness of pains reduces to representation of *H* on functional/teleological grounds, all the explanatory work that is being done by pain's affect is done by these forward-looking downstream consumer oriented causal connections that constitute pain states' functional role. Therefore, any explanatory appeal to pain's affect representing *H* is redundant. There is nothing that pain's affect's representing *H* can explain that cannot be explained by these pain states' historical and present functional role. If so, representationalism about pain affect cuts no explanatory ice.

Consider a psychofunctionalist (or, a functionalist — see Block 1980 for the distinction) about pain *affect* that says that the painfulness of pain consists in nothing but the states of a certain type *playing* a certain psychofunctional role *PF*.³³ The state type in question is one whose tokens indicate (or purport to indicate) certain types of bodily disturbance — let's call the

³² Actually, this is almost accepted by C&T. See their fn. 9 on p. 107 of their (2011).

³³ For such accounts, see Author 2000, 2005; Clark 2005. Please note that the psychofunctionalism entertained here is only about *affect*. Some form of intentionalism about the perceptual or sensory-discriminative phenomenology of experiences is not only consistent with such a psychofunctionalist account but in fact quite naturally invites it (see Authors (in prep.)).

state type *S*. The psychofunctional role may include all sorts of information processing, biasing or predisposing the possessors of these states to behave in certain ways, causing other mental states, reordering preferences and (pre)motor priorities, setting certain range of motivational parameters, etc. According to such a psychofunctionalist, an evolutionary explanation may be apt for why tokens of *S* came to typically exhibit *PF*. In fact, she can give pretty much the same teleological story as that of C&T's about how certain teleological systems capable of detecting certain physical disturbances have acquired state type *S* with *PF*. We can ask: by refusing to make the additional claim that tokens of *S* thereby represent *H*, what explanation is this psychofunctionalist missing? We can see none.

On C&T's view, the painfulness of pain experiences represents harmfulness because the pain experiences have played and still play a certain psychofunctional role. It isn't the case that they play this role because their painfulness represents harm. Sensory representations are useful because of what they represent. They are used and processed to extract information about what they represent. This is to say that they are usually subject to causal/functional processing *because* of what they represent. They are also the epistemic and causal sources of our sensory and perceptual *concepts* that apply to what these sensory representations represent. But on C&T's account, that the painfulness represents harm due to its historical/functional role doesn't (in fact, can't) figure in these sorts of explanations adverting to representational content. Consider, in general, the claim that the pleasant and unpleasant affect represents objects of perceptual experiences as beneficial or harmful and the representationalist account of this fact given by C&T. Are these experiences the epistemic and causal sources of our concepts BENEFICIAL, HARMFUL, in the way, say, visual experiences of red are the epistemic and causal sources of our concept RED? No. Do the pleasantness or unpleasantness of experiences *teach* us (on the basis of their putative representational content) that their objects are beneficial or harmful in the way experiences of red do about the redness of their objects? No. So not only adding the representationalist claim to psychofunctionalism about affect doesn't do any explanatory work, but it also results in wrong predictions. (Indeed, this was one of the main argumentative themes against SR throughout this paper — especially if SR is claimed to entail OV.) This is not to argue that psychofunctionalism about pain affect (or affect, in general) is true, but it is to argue that it is superior to representationalism about (pain) affect representing *H* on teleosemantic/functional grounds.

There are of course more general considerations against SR about affect. The unpleasantness or pleasantness of an experience, on SR, is just the property of this experience's representing its object as harmful or beneficial. Clearly these properties have different second-order properties. For instance, affective experiences such as pain are *intrinsic motivators* — they motivate the experiencer solely on the basis of their presence. This is not true at all about the property of an experience's representing its object as harmful or beneficial. C&T may appeal to the functional role of these experiences, but this would be an appeal to the wrong property. Functional role can explain an experience's being an intrinsic motivator only qua causal role, not qua representor of harm/benefit. Moreover, making the further claim that that functional role

also constitutes the experience's representing harm or benefit doesn't add anything to the explanation by functional role alone. It's redundant at best.

There is also the fact that affect attaches to experiences in all sensory modalities that give rise to conscious experiences. A representationalist needs to find a common second-order property for *all* pleasant or unpleasant experiences to represent across all modalities. It is very doubtful whether she has the resources to do that on the basis of what has historically harmed or benefited organisms.

Further, affective responses vary enormously across as well as within individuals with very strong conative and behavioral effects even for *subtle* differences in the representational content of experiences. For instance, while I like dark chocolate, I dislike milk chocolate; while I like cauliflower, I hate broccoli. But my brother's preferences are just the other way around: He likes milk chocolate, but hates dark chocolate, etc. Further, my preferences were just like my brother's when we were growing up. Tons of variation! Note that if the representationalist sticks to the second-order properties like apt to harm or benefit, misrepresentation will *be* the rule (not the exception!) in modern history when explaining affective phenomenology — small wonder then why the truth-values of our first-order affective *judgments* don't follow those of our affective *experiences* (even supposing that experiences represent harm or benefit). These properties are just too coarse grained to do justice to our intuitions about affective aspects of our experiences. Psychofunctionalism doesn't suffer from these problems.³⁴

8 Conclusion

We have argued (in §§1–5) that representationalism about phenomenal character is false on the ground that representationalists can't give a satisfactory account of affective qualities consistent with the core tenets of their representationalism. We have then argued (in §6) that the kind of psychosemantics that Cutter and Tye (2011) give for affective qualities, contrary to their own advertisement, is not a tracking (informational) but a teleological psychosemantics with its own set of problems for affect. In §7, we have claimed that the basic machinery for running a teleological and consumer-oriented psychosemantics for affective qualities is equally available

³⁴ Brendan O'Sullivan and Robert Schroer (2012) argue against functionalism about affect. The gist of their argument is roughly this. Among the typical causes of an affective experience are conative attitudes like desiring and liking the experience. If functionalism says that affective character of an experience is (partly) constituted by such causal connections, then it gets the explanatory order wrong: one's experience doesn't feel good (partly) because one likes it or desires it — the explanation should go in the other direction. But if it goes in the other direction, functionalism doesn't explain affect: it just assumes it. We are not sure this is really a problem for our kind of functionalism about affect. Our functionalism is not analytic functionalism or a functionalism whose terms are drawn exclusively from folk psychology or personal level cognitive psychology. We can't give a proper defense of psychofunctionalism here — but see Author (2000: 559-60) and Clark (2005) to see how the sort of problem O'Sullivan and Schroer raise can be avoided. They, along with Bain (2012), also make the charge that a non-representationalist about affect cannot do justice to the rationalizing role of affect. We think that a psychofunctionalist can in fact do a better job than a representationalist about this issue — see our (in prep.).

for defending psychofunctionalism about affect without tangling oneself with the extravagances of (strong) representationalism. We are physicalist/naturalist about experiential phenomenology. But we don't think SR is the right way to naturalize it. As we have shown, this comes out particularly clearly and loudly in the case of affective phenomenology. Many physicalists and reductionists about phenomenal consciousness think that SR is false without thinking that intentionalism about (most or all) phenomenology is false.³⁵ The point of discussion in §§6–7 was therefore to show that some versions of naturalistic (teleological or psychofunctionalist) psychosemantics may be available, not for *reducing* affective phenomenal content to representational content, but for showing how affective qualities can nevertheless be intentional — if a physicalist/naturalist would like to maintain intentionalism about affective phenomenology without holding SR. Partly for reasons expressed in the last section, we ourselves think that psychofunctionalism is sufficient for the metaphysics of affective qualities and that affective qualities (as instantiated in experiences) don't have intentional content. But perhaps a non-SR naturalist about phenomenal consciousness can make a case for affective qualities' having intentional content and why this is a good thing.

We conclude that the headache caused by affect is fatal for representationalists.³⁶

References

- Armstrong, D. M. (1962). *Bodily Sensations*. London, Routledge and Kegan Paul.
- Armstrong, D. M. (1968). *A Materialist Theory of the Mind*. New York, Humanities Press.
- Aydede, M. (2000). "An Analysis of Pleasure Vis-à-Vis Pain," *Philosophy and Phenomenological Research*, 61(3): 537–570.
- Aydede, M. (2001). "Naturalism, introspection, and direct realism about pain," *Consciousness and Emotion*, 2(1): 29–73.
- Aydede, M. (2002). "Is Introspection Inferential?" In *Privileged Access: Philosophical Accounts of Self-Knowledge*, Brie Gertler (Editor), Ashgate Publishing (*Epistemology and Mind Series*)
- Aydede, M. (ed.), (2006). "The Main Difficulty with Pain." In *Pain: New Essays on Its Nature and the Methodology of Its Study*, M. Aydede (ed.), Cambridge, Mass.: MIT Press.
- Aydede, M. (2009). "Is Feeling Pain the Perception of Something?," *Journal of Philosophy*, 106(10): 531–67.
- Aydede, M. (2010). "Pain", *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2010/entries/pain/>>

³⁵ We are not alone in thinking this. See, for instance, Perkins (1983), Lycan (1987), Clark (1993, 2000), Shoemaker (1996), Perry (2001), Loar (2003), Matthen (2005), Author (XXXX) along with many, many others.

³⁶ For useful feedback and commentary, many thanks to David Bain, Jonathan Cohen, Jennifer Corns, Adam Pautz.

- Aydede, M. (forthcoming). "Is the Experience of Pain Transparent?" forthcoming in *Review of Philosophy and Psychology*.
- Aydede, M. and Güzeldere, G. (2005). "Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness," *Noûs*, 39(2): 197–255.
- Bain, D. (2003). "Intentionalism and Pain," *Philosophical Quarterly*, 53(213): 502–522.
- Bain, D. (2007). "The Location of Pains," *Philosophical Papers*, 36(2): 171–205.
- Bain, D. (2012). "What Makes Pains Unpleasant?" forthcoming in *Philosophical Studies*.
- Block, N. (1980). "Troubles with Functionalism." In *Readings in Philosophy of Psychology* edited by N. Block. Cambridge, MA, Harvard University Press.
- Byrne, A. (2009). "Experience and content," *The Philosophical Quarterly*, 59(236): 429–51.
- Byrne, A. (2012). "Knowing What I See." In *Introspection and Consciousness*, ed. D. Smithies and D. Stoljar, New York: Oxford University Press.
- Byrne, A., & Hilbert, D. R. (1997). "Colors and Reflectances." In *Readings on Color (Vol. 1)*, Alex Byrne & David Hilbert (Eds.), Cambridge, Massachusetts: MIT Press.
- Byrne, A., & Hilbert, D. (2003). "Color realism and color science," *Behavioral and Brain Sciences*, 26(01): 3–21.
- Byrne, A., & Tye, M. (2006). "Qualia ain't in the head," *Nous*, 40(2): 241–255.
- Clark, A. (1993). *Sensory Qualities*. Oxford, New York, Oxford University Press.
- Clark, A. (2000). *A Theory of Sentience*. New York: Oxford.
- Clark, A. (2005). "Painfulness is Not a Quale." In *Pain: New Essays on Its Nature and the Methodology of Its Study*. M. Aydede, Ed., Cambridge, Massachusetts: MIT Press.: 177-198.
- Cutter, B., & Tye, M. (2011). "Tracking Representationalism and the Painfulness of Pain," *Philosophical Issues*, 21(1): 90–109.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, Massachusetts, MIT Press.
- Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Dretske, F. (1999). "The Mind's Awareness of Itself." *Philosophical Studies*, 95(1-2): 103–124.
- Dretske, F. (2003). "How Do You Know You Are Not A Zombie?" In *Privileged Access: Philosophical Accounts of Self-Knowledge*. B. Gertler. Hampshire, UK, Ashgate Publishing.
- Dretske, F. (2006). "The Epistemology of Pain." In *Pain: New Essays on Its Nature and the Methodology of Its Study*. M. Aydede, Ed., Cambridge, Mass.: MIT Press.
- Dretske, F. (2012). "Awareness and Authority: Skeptical Doubts about Self-Knowledge." In *Introspection and Consciousness*, ed. D. Smithies and D. Stoljar, New York: Oxford University Press.

- Harman, G. (1990). "The Intrinsic Quality of Experience." *Philosophical Perspectives* 4: 31–52.
- Kripke, S. (1979). "A Puzzle about Belief". In *Meaning and Use*, edited by A. Margalit. Dordrecht and Boston: Reidel.
- Lycan, W. G. (1987). *Consciousness*. Cambridge, Massachusetts, MIT Press.
- Lycan, W. G. (1996). *Consciousness and Experience*. Cambridge, MA: MIT.
- Lycan, W. G. (2008). "Representational Theories of Consciousness", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/consciousness-representational/>>.
- O'Sullivan, B., & Schroer, R. (2012). "Painful Reasons: Representationalism as a Theory of Pain," *The Philosophical Quarterly* 62: 737-758.
- Pitcher, G. (1970). "Pain Perception," *The Philosophical Review*, 79(3): 368–393.
- Robinson, W. S. (2006). "What Is it Like to Like?," *Philosophical Psychology*, 19(6): 743–765.
- Seager, W. (2002). "Emotional Introspection," *Consciousness and Cognition*, 11(4): 666–687.
- Tye, M. (1995). *Ten Problems of Consciousness: A representational theory of the phenomenal mind*. Cambridge: MIT.
- Tye, M. (1996). "Orgasms Again," *Philosophical Issues*, 7: 51-54.
- Tye, M. (1997). "A Representational Theory of Pains and their Phenomenal Character." In *The Nature of Consciousness: Philosophical Debates*. N. Block, O. Flanagan and G. Güzeldere, Eds. Cambridge, Massachusetts: MIT Press.
- Tye, M. (2000). *Consciousness, Color, and Content*, Cambridge, Mass: The MIT Press.
- Tye, M. (2002). "To PANIC or Not to PANIC? — Reply to Byrne." Retrieved in October 2012 from http://host.uniroma3.it/progetti/kant/field/tyesymp_replytobyrne.htm
- Tye, M. (2005a). "Another Look at Representationalism about Pain." In *Pain: New Essays on Its Nature and the Methodology of Its Study*. M. Aydede, Ed., Cambridge, Massachusetts: MIT Press.
- Tye, M. (2005b). "In Defense of Representationalism: Reply to Commentaries." In *Pain: New Essays on Its Nature and the Methodology of Its Study*. M. Aydede, Ed., Cambridge, Massachusetts: MIT Press.
- Tye, M. (2009). *Consciousness Revisited: Materialism without Phenomenal Concepts*. Cambridge, Massachusetts: MIT Press.