

Reductive Representationalism and Emotional Phenomenology

Uriah Kriegel

Forthcoming in *Midwest Studies in Philosophy*

Introduction/Abstract

A prominent view of phenomenal consciousness combines two claims: (i) the identity conditions of phenomenally conscious states can be fully accounted for in terms of these states' representational content; (ii) this representational content can be fully accounted for in non-phenomenal terms. This paper presents an argument against this view. The core idea is that the identity conditions of phenomenally conscious states are not fixed entirely by *what* these states represent (their representational contents), but depend in part on *how* they represent (their representational attitudes or modes). The argument highlights the myriad liabilities and difficulties one must accrue if she tries to appeal *only* to what phenomenally conscious states represent in accounting for their phenomenal individuation.

1. Reductive Representationalism

One of the central research areas in the philosophy of mind of the past quarter-century has been the attempt to provide a reductive philosophical explanation of phenomenal consciousness. In contrast to a *scientific* explanation of a phenomenon, which often states the causal conditions for the phenomenon's occurrence, a *philosophical* explanation states rather the identity and existence conditions of the phenomenon. It attempts to answer the following question:

(Q) What makes a phenomenally conscious state S (i) the phenomenally conscious state it is and (ii) a phenomenally conscious state at all?

The answer to (i) states the identity conditions of phenomenally conscious states, the answer the (ii) their existence conditions.

A *reductive* philosophical explanation of phenomenal consciousness is any account that answers Q in exclusively non-phenomenal terms. Several such have been offered over the past quarter-century. It is fair to say, however, that the most

dominant has been representationalism (sometimes also called ‘intentionalism’ or ‘first-order representationalism’).¹ Representationalism admits of various formulations. Originally, it was often framed in terms of supervenience: phenomenal character metaphysically supervenes on representational content. Here ‘phenomenal character’ stands for the totality of phenomenal properties instantiated by a conscious state and ‘representational content’ for what that state represents. The idea, then, was that there is no pair of metaphysically possible worlds in which the totality of phenomenal properties instantiated by some conscious state differs but what the state represents is the same. It is unclear, however, that ‘supervenience representationalism’ (if you will) really provides an answer to Q. For it does not *identify* what makes a phenomenally conscious state the phenomenally conscious state it is and a phenomenally conscious state at all; it only tells us that, whatever that is, it cannot vary independently of representational content.

Since the inception of representationalism, philosophy has witnessed the emergence of various ‘deeper’ relations than supervenience, such as superdupervenience (Horgan 1993) and especially grounding (Fine 2001).² We may think of ‘grounding representationalism’ as the following thesis: phenomenal character is grounded in representational content, in the sense that a conscious state instantiates the totality of its phenomenal properties *in virtue of* representing what it represents. The ‘in virtue of’ here is supposed to be explanatory: when *p* holds in virtue of *q*, *q* *accounts* for the fact that *p* holds. Thus grounding representationalism at least has the *potential* to answer Q.

Rather than seeking the best specific notions in terms of which to frame an account of phenomenal consciousness that delivers an answer to Q, though, we might simply answer Q *directly* in terms of representational content. As a first pass, we might try:

(FIRST) What makes a phenomenally conscious state *S* the phenomenally conscious state it is, and a phenomenally conscious state at all, is the fact that *S* has the representational content it does, and has one at all.

The problem here is that many non-conscious mental states have representational contents. So just having a representational content cannot be what makes a mental state conscious. At a minimum, there must be a special *kind* of representational content, such that having a content of *that* kind makes a mental state phenomenally conscious. A second pass might therefore be:

(SECOND) There is a kind of representational content *C*, such that for every phenomenally conscious state *S*, what makes *S* the phenomenally conscious

state it is, and a phenomenally conscious state at all, is that S has the C-content it does, and has a C-content at all.

As noted, a state's representational content is a matter of *what the state represents*. So, we may elucidate the relevant *kind* of content in terms of the kind of *entity* represented. This leads to the following third pass:

(THIRD) There is a class of entities E, such that for every phenomenally conscious state S, what makes S the phenomenally conscious state it is, and a phenomenally conscious state at all, is that S represents the members of E it does, and represents members of E at all.

Two clarifications are in order. First, 'entity' is used as the most generic ontological term, covering not only concrete particulars but also properties, property-instances, events, states of affairs and whatever else one allows in one's ontology; it is the ontological *summum genus*. Secondly, the verb 'represents' and the locution 'what is represented' can be heard in two importantly different ways. Consider a hallucination of a yellow lemon in a lemonless world. There is a way of hearing representation talk where what is represented by the hallucination is a yellow lemon, and another way where *nothing* is represented. It is the first hearing of representation talk that should be adopted here.

In the bulk of this paper, I use the term 'representationalism' essentially as a label for THIRD. (I will address the question of whether representationalism may admit of more permissive formulations in §2.) In offering a response to Q, THIRD already provides a philosophical explanation of phenomenal consciousness. To provide a *reductive* philosophical explanation, however, THIRD must be conjoined with another claim, namely, that representation can be accounted for in entirely non-phenomenal terms. The claim is that it is possible to say what makes a conscious state represent what it does (and represent at all) without invoking any phenomenal notion. In the 'material mode of speech': there is a kind of non-phenomenal fact – that is, a fact without phenomenal constituents – that makes conscious states represent what they do (and at all). Typically, representationalists cite facts about the causal, informational, and/or teleological relations conscious states bear to worldly entities.³ But the minimal requisite claim is simply:

(REDUCTIVE) There is a class of facts F, such that (i) no member of F has a phenomenal constituent and (ii) for every phenomenally conscious state S, what makes S represent the entities it does, and represent entities at all, are members of F.

I take the conjunction of THIRD and REDUCTIVE to constitute the core of representationalism understood as an attempt at a reductive philosophical explanation of phenomenal consciousness. Indeed, from now on I will simply use 'reductive representationalism' as a label for the following synthesis of THIRD and REDUCTIVE:

(REDUCTIVE REPRESENTATIONALISM) There is a class of entities E and a type of relation R, such that (i) R is a non-phenomenal relation and (ii) for every phenomenally conscious state S, what makes S the phenomenally conscious state it is, and a phenomenally conscious state at all, is that S bears R to the members of E it does, and bears R to members of E at all.

Arguably, something like this has animated the dominant approach to the reductive philosophical explanation of phenomenal consciousness over the past quarter-century. I now turn to argue against it.

2. Attitudinal-Representational Properties

Early arguments against representationalism have by and large been arguments by counter-examples (e.g., Block 1990). Among those proffered by anti-representationalists were cases focusing on alleged representation of a single environmental feature in two different perceptual modalities. Block (1996), for example, claimed that a visual and an auditory experience can have the same representational content, but vary in phenomenal character. Imagine a fly buzzing just over your head, fast enough that you do not quite realize it is a fly, but still have a visual impression as of *something overhead*. Suppose in addition you also *hear* that 'something' overhead. Both experiences have the content (we would express as) 'something overhead' (or perhaps 'there is something overhead'); but while one experience carries that content *visually*, the other carries it *auditorily*. The two experiences thus deploy phenomenally different *modes* of representation, even as *what* they represent is one and the same. In a similar vein, Lopes (2000) claimed that some atypical subjects may recruit certain modalities to represent environmental features typically represented by other modalities. In particular, some blind people develop the capacity to detect the shapes of objects by hearing, suggesting that they have auditory experiences of shape where we only have visual experiences of shape. Their shape experiences and ours both represent shape, but ours represent shape visually while theirs represent it auditorily. The point is that visual and auditory experiences are always *different* phenomenally conscious states, so if they bear the same (e.g., teleo-informational) relations to the same entities (say, the property of rectangularity), bearing that relation to those entities cannot

be what makes them the conscious states they are – contrary to reductive representationalism. Call this the *argument from perceptual modes*.

Various responses to these prima facie counter-examples are possible, but the most natural ones appeal to special environmental features which lend themselves to representation in only one modality. On this line, there is a visible-only feature of rectangles, which we may call ‘the look of rectangularity,’ as well as an audible-only feature of rectangles, which we may call ‘the sound of rectangularity.’ Our visual experiences do not in fact represent rectangularity, but rather the-look-of-rectangularity, while the aforementioned atypical auditory experiences do not actually represent rectangularity, but rather the-sound-of-rectangularity. A more flexible variant of the same line might claim that our visual experiences do represent rectangularity, but do so *by* representing the-look-of-rectangularity – while the relevant auditory experiences represent rectangularity by representing the-sound-of-rectangularity.⁴ Either way, the visual and auditory experiences turn out to differ not only in phenomenal character, but also in representational content – contrary to what Block and Lopes claim.⁵

What I want to develop in the bulk of this paper is an argument against reductive representationalism that is similar in structure to the perceptual-mode argument, but where the move to special environmental features (in the style of the-look-of-rectangularity and the-sound-of-rectangularity) is much less plausible. To do so, I will focus on non-perceptual experiences, notably *emotional* experiences. In the remainder of this section, I present the back-of-the-envelope version of the argument. I will develop the argument more fully in the next section.

The argument focuses on the kind of emotional state we tend to report using non-propositional, ‘objectual’ constructions such as ‘I am afraid of snakes,’ ‘I love you,’ and ‘I am disappointed with the loss,’ rather than those we tend to report using propositional constructions, such as ‘I am afraid that there will be snakes there,’ ‘I love that you always listen,’ and ‘I am disappointed that the team did not make a better effort.’ I will address the objection that the non-propositional constructions are in fact misleading, and that in truth all emotional states have propositional content, toward the end of §3. Until then, I assume that there really are distinctive objectual, non-propositional emotions in our psychological repertoire.



Consider two friends coming across a dog. One of them feels affection toward the dog, whereas the other is afraid of the dog. On the face of it, the affection experience and the fear experience – both phenomenally conscious states – are about one and the same thing: the dog. Nonetheless, they are *different* phenomenally conscious

states – what it is like to feel afraid of the dog is very unlike what it is like to feel affection for the dog. Likewise, imagine you feel admiration toward your spouse at one moment, and this prompts you the next moment to vividly feel your love for him or her. Your experience of admiration and your experience of love feel rather different, but what they are about is the same thing – your spouse.

To handle such cases with the same strategy as used by for perceptual modalities, the reductive representationalist would need to make two claims: (1) there are special properties which can only be admired, loved, feared, and so on; (2) these properties appear in the representational contents of the phenomenally conscious states just described. On first glance, this might seem reasonably plausible. It is traditionally thought that emotions have ‘formal objects’: fear is directed *at the dangerous*, admiration is directed *at the admirable*, love is in some sense directed *at the lovely*, affection in some sense *at the affecting*. It might be suggested that one of the two friends represents the dog as dangerous, while the other represents the dog as affecting; and that while you represent your spouse one moment as admirable, the next moment you represent him or her as lovely. Thus there are clear differences in representational content between the relevant pairs of experience, differences which match the evident differences in phenomenal character.

Upon closer inspection, however, it is implausible to suppose that the relevant special properties show up in the emotional experiences’ *contents*. On the face of it, one fears *a dog*, not the dog’s dangerousness, just as one admires *one’s spouse*, not the spouse’s admirability. It would be very strange indeed to admire a person’s admirability rather than the person herself, and it would be likewise strange to fear a thing’s dangerousness rather than the thing itself. It is true, of course, that fear of a dog involves in some sense a certain mental commitment to the dog’s dangerousness. But that commitment seems built into the distinctive *attitude* of fearing, rather than showing up in the *content* of the fear. Arguably, what fear *is*, as a type of mental state, is a commitment to the feared object’s dangerousness – a *felt* commitment in the case of conscious experiences of fear. In other words: danger is not part of *what* fear represents, it is an aspect of *how* fear represents. We may summarize the basic point with the following formulation:

(FEAR) The experience of fearing *x* involves representing-as-dangerous *x* rather than the representing *x*-as-dangerous.

The idea is that commitment to dangerousness shows up in fear as a modification of the *manner* of representation, rather than as an element among others represented by the fear. The same point can be made about admiration:

(ADMIRATION) The experience of admiring x involves representing-as-admirable x rather than the representing x -as-admirable.

Thus admirability is not part of *what* admiration represents; rather, it is an aspect of *how* admiration represents (whatever it does).⁶

Call such properties as representing-as-dangerous and representing-as-admirable *attitudinal-representational properties*. They are representational properties in the sense that they cast worldly entities a certain way. A mental state that were entirely non-representational – represented nothing – could not instantiate such properties.⁷ At the same time, these are representational properties that a mental state does not instantiate in virtue of its representational content. Instead, it instantiates them in virtue of the distinctive *attitude*, or *mode*, it deploys. They do not have to do with what the state represents, but with how it represents. This paper's positive thesis is that phenomenally conscious states *have* attitudinal-representational properties and, moreover, phenomenally individuate sensitively to the attitudinal-representational properties they have. More precisely:

(POSITIVE) For some phenomenally conscious state S , there is an attitudinal-representational property A , such that *part* of what makes S the phenomenally conscious state it is is that S instantiates A .⁸

POSITIVE is inconsistent with reductive representationalism, indeed with representationalism itself, as formulated in THIRD. THIRD asserts that what makes a phenomenally conscious state the phenomenally conscious state it is is just its representational content, what it represents. But POSITIVE denies this, claiming that part of what makes that state the phenomenally conscious state it is has nothing to do with *what* the state represents; it comes rather from *how* it represents, the attitudinal-representational properties it exhibits. Call this the *argument from non-perceptual modes*.

(Note that POSITIVE does not comment on what makes a state a phenomenally conscious state *at all* – only on what makes it the phenomenally conscious state it is. That is, it concerns only the identity conditions of phenomenally conscious states, not their existence conditions. There may well be reasons to suspect that the existence condition of phenomenality cannot be captured by representational content either, since on the face of it, any entity which can be represented consciously can also be represented non-consciously – there are no entities which lend themselves only to conscious representation (Kriegel 2002, Chalmers 2004). But the argument from non-perceptual modes attempts to show that even if the representationalist retreats to a more circumscribed claim about identity conditions only, she faces insurmountable difficulties.⁹)



Before developing the argument from non-perceptual modes in more detail, let us consider how the reductive representationalist could weaken her thesis so as to be made consistent with POSITIVE. As just noted, attitudinal-representational properties are in an important sense *representational* properties. For they are, after all, properties of *representing-as-F*. In §1, we have formulated representationalism as a thesis about representational *contents*: as the claim that phenomenally conscious states' identity and existence conditions are given by *what* they represent. A weaker formulation, however, might appeal to representational *properties* in a sufficiently generic sense that covers attitudinal-representational properties.¹⁰

It might be useful here to draw a distinction between *pure* and *impure* representationalism (Chalmers 2004).¹¹ Pure representationalism individuates phenomenally conscious states exclusively in terms of those states' representational *contents* (what they represent). Impure representationalism individuates them partially in terms of their of *attitudinal-representational* properties (how they represent).

Impure representationalism can certainly be found in modern philosophy of mind (Crane 2003, 2009, Chalmers 2004). But most reductive representationalists have been purists. Fred Dretske led the way:

An experience of movement – whether it be visual, tactile, or kinesthetic – has its qualitative character defined by what it is an experience (representation) of, and if these experiences are all of the same property, they are, subjectively, with respect to this single property, the same kind of experience. Even when the senses of overlap in their representational efforts – as they do in the case of spatial properties – they ... represent different ranges of determinable properties. (Dretske 1995: 94-5)

Michael Tye drew the conclusion:

Phenomenal character (or what it is like) is one and the same as a certain sort of intentional content. (Tye 1995: 137)

Alex Byrne is usefully explicit on all this:

... all intentionalists [or representatioanalists] agree that *within* a (paradigmatic) perceptual modality, if two possible experiences differ in phenomenal character, they differ in content. *Intermodal* intentionalists hold, while *intramodal* intentionalists deny, that the phenomenal difference *between* perceptual modalities – between visual and auditory experiences, for example – is determined by a difference in content... [T]he conclusion of the main argument to follow is this: intermodal unrestricted intentionalism. (Byrne 2001: 205-6; italics original)

As is Michael Thau:

[Some philosophical problems may] provide arguments that facts about *what* is represented are insufficient to adequately distinguish perceptual experiences from each other, beliefs from each other, and perceptual experiences from beliefs. And, hence, they suggest that facts about *how* perceptual experiences and beliefs represents are required to properly individuate them... My aim in this book is to bring the view out into the open and, in doing so, show that it is mistaken. (Thau 2002: 14; italics original)

My target here is the pure reductive representationalist like Dretske, Tye, Byrne, and Thau. The central negative thesis is accordingly:

(NEGATIVE) POSITIVE undercuts REDUCTIVE REPRESENTATIONALISM.

Together, POSITIVE and NEGATIVE constitute this paper's take-home conclusion.¹²

3. The Argument from Non-Perceptual Modes

Consider again the person who feels affection for a dog and the person who feels afraid of the same dog. On the face of it, their experiences differ phenomenally but share representational content. There are two general options for denying that this is in fact the case. One is to deny that the two experiences are phenomenally different. This seems completely desperate – it seems clearly possible for us to tell whether we are afraid or feel affectionate by noting what it is like for us.¹³ The other option is to deny that the fear and affection experiences in fact have the exact same representational content. My goal in this section is to show just how difficult and costly it would be to try to pursue this approach.



The simplest strategy for pursuing this approach is simply to deny that we actually just fear a dog and love a person. In fact, it might be said, we fear the dog's dangerousness and love the person's loveliness. It has been pointed out above that this is not at all how we ordinarily *speak* of our emotional experiences. We speak as though we love persons and fear dogs. But the representationalist may dismiss this as a superficial feature of language. The psychological reality, she may insist, is that we fear the dog's dangerousness, not just the dog.

The first thing to note about this suggestion is that it appears to render all fears irrational. We can make sense of a person who fears a dog, because dogs sometimes bite. But unlike dogs, the property of dangerousness does not bite, and so it would be completely irrational for the person to fear it. In fearing the dog's dangerousness, she would be committing a sort of category mistake: what she fears cannot harm her.

Furthermore, it is a platitude about fear that it tends to induce a fight-or-flight response. Again, we can understand why someone who is afraid of a dog would fight that dog or flee him. But it is impossible to fight a property, and trying to flee it would be senseless. The point is that we tend to fight or flee that which we fear, so if what we tend to fight or flee is the dog himself, rather than his dangerousness, then it must be the dog himself that we fear, not his dangerousness. Likewise, a person overcome with affection for her child may reasonably wish to embrace the object of her affection. But embracing a property is neither satisfying nor feasible.

The representationalist may suggest that what we fear, on her view, is not simply the dog's property of being dangerous, but the *state of affairs* of the dog being dangerous, a state of affairs that after all has a concrete, spatiotemporal constituent in the dog. However, it is not immediately clear how this helps: states of affairs do not bite any more than properties do, and trying to fight them would be just as quixotic. The representationalist may point out that although the relevant state of affairs cannot bite, it has a constituent that can, a constituent which may moreover be fought or fled. But this would only rationalize the part of the fear that is concerned with that part of the state of affairs. The additional constituent of the state of affairs – the property of dangerousness – would seem to be something of a dangler.

A much more natural claim is that we fear the dog *because* of his dangerousness (or *in virtue of* his being dangerous). Here the dog's dangerousness is not part of *what* is feared, however, but functions rather as the *ground* or *reason* for fear. It captures the *appropriateness conditions* of the fear. If the dog really does pose a danger, one's fear experience is appropriate or fitting; if it poses no danger, it is not. Indeed, we may ask ourselves: *why* does danger capture the appropriateness conditions of fear? To this question, the proponent of attitudinal-representational properties has a straightforward answer: danger captures the appropriateness conditions of fear because fear is essentially the mental state which represents-as-dangerous. Likewise, admirability captures the appropriateness conditions of admiration because admiration is essentially the mental state which represents-as-admirable. Compare: belief has *truth* conditions because it is essentially the state which represents-as-true. It is each mental state's distinctive attitudinal-representational property that sets the kind of appropriateness conditions that apply to that state.

It should be remarked that the case of fear and danger is in a certain way the most hospitable to the representationalist's agenda. For we have a clear grasp of the nature of danger that is independent of our experience of fear. This is not always the

case. We may say that love represents the beloved *as lovely*. But what does 'lovely' really stand for in this context? It seems to be used in a completely technical sense. (Mike Tyson's partner may love him without any illusions about how 'lovely' he is *in the ordinary sense of the word*.) But do we have any grasp of what the property of loveliness denoted in this technical use of 'lovely' exactly is? Insofar as we have any grasp of the relevant property, it seems entirely derived from our grasp of love as a particular type of experience. To that extent, it is unclear how we are supposed to evaluate the claim the experience of love represents loveliness. Virtually identical remarks can be made about the property of 'being affecting' that felt affection is claimed to represent (and for that matter the property of being disappointing that the experience of disappointment would have to represent, the property of being angering that anger would have to represent, and so on).



There is, in any case, a deeper problem with the suggestion under consideration. Recall that in response to the argument from perceptual modes, the representationalist posited visible-only and audible-only properties that visual and auditory experiences were claimed to represent. To extend the strategy to emotional experiences, the representationalist would have to posit fearable-only properties, loveable-only properties, and so on. Yet danger and loveliness are not such properties. We can also *believe* that a dog is dangerous, *consider* whether he is dangerous, and be in many other mental states that involve the representation of dangerousness. The proponent of attitudinal-representational properties has a straightforward explanation of the difference between fearing a dog and believing that he is dangerous: the former involves representing-as-dangerous the dog, the latter representing the-dog-as-dangerous. But a representationalist who hopes to capture the phenomenal character of fear in terms of the property of dangerousness being represented owes us some other story about the difference between the two.

The only live option for the representationalist seems to be to posit an extra-special property – more special than simple dangerousness – that could only be feared. In the perceptual case, the representationalist posited such properties as the look-of-rectangularity and the sound-of-rectangularity. It would seem that here too, she must posit something like the fearsomeness-of-danger. The fearsomeness-of-danger is a property that tends to co-occur with dangerousness, but is nonetheless different – as different as the look-of-rectangularity is from rectangularity. Similarly exotic properties would have to be posited for all other emotional experiences. In some cases, there would be remotely reasonable-sounding names for those properties, such as the 'loveableness-of-loveliness'; but for such experiences as admiration, frustration, and disappointment there would be none elegant. All the

same, such extra-special properties must be recognized in each case (since in each case a corresponding belief can be had that would target the original property).

One may wonder whether there really are such properties, and indeed whether we fully understand what these properties are supposed to be. But the deeper problem is different. The purpose of introducing visible-only and audible-only properties was to rid us of vision and audition as distinctive *modes* of perception. The new picture was to replace the variety of perceptual *ways* of representing the same entities with a single generic way of representing *different* entities. This generic way of representing is a sort of *sensory entertaining*, which can take as objects a variety of very different entities, including visible-only properties, audible-only properties, and so on. On this view, *seeing* a shape is not a matter of visually-representing the shape, but a matter of entertaining the look-of-the-shape; hearing something overhead is not a matter of auditorily-representing overhead-ness, but of entertaining the-sound-of-overhead-ness. There is no point in introducing the visible-only properties, audible-only properties, olfactible-only properties etc. if we *also* preserve distinct perceptual *modes* of representing these properties. That would be a kind of theoretical overkill, where one of the two posits (special perceptible properties or perceptual modes) is bound to be explanatorily preempted by the other.

Accordingly, extending the same strategy to non-perceptual experiences, by introducing such properties as the fearsomeness-of-danger and the loveableness-of-loveliness, would involve ridding ourselves of fear, love, and other emotional modes in favor of a single, uniform mode of *emotional entertaining*. To fear a dog, on this view, is to entertain the dog's fearsome-dangerousness; to love a person is to entertain the person's loveable-loveliness; and so on. All emotions employ a single, uniform mode of being aware of something – they differ only in what one is aware of under that mode.

In fact, given that the representationalist ultimately wants to account also for the phenomenal difference between seeing a dog and fearing a dog, she cannot even avail herself of a distinction between sensory entertaining and emotional entertaining. Rather, there must be a single kind of *pure entertaining*, which can be aimed at sensible properties or at emotable properties; the difference between seeing a dog and fearing a dog is thus ultimately a matter of what one entertains in this pure way. Thus to fulfill its promise, (reductive) representationalism must hold that *all* phenomenally conscious states share this single way of aiming at external entities.

We can see how implausible this is by comparing and contrasting the case of belief and desire. There are of course several differences between belief and desire, including in functional role and ‘direction of fit.’ But it is often thought that there is also a fundamental difference in the *way* belief and desire represent what they do. In particular, it is natural to claim that belief that *p* involves a commitment to the *truth* of *p* (or to the *obtaining* of *p*), whereas a desire that *p* involves a commitment to the *goodness* of *p* (in some suitably generic sense of ‘goodness’). Believing that there is whiskey at the party in some sense commits one to the truth of the proposition < there is whisky at the party > (or to the obtaining of the state of affairs of there being whisky at the party). In contrast, desiring that there be whisky at the party commits one to the goodness of the state of affairs of there being whiskey at the party. However, we do not tend to think that believing that *p* is just entertaining that *p* is true (or that *p* obtains), nor that desiring that *p* is just entertaining that *p* is (or would be) good. Infants and animals lacking the concepts of truth and goodness can still believe and desire, which suggests that the concepts of truth and goodness are *not* constituents of their relevant mental state’s propositional contents. More deeply, it seems to be in the nature of entertaining that as long as one merely entertains the truth of *p*, one is not yet *committed* to the truth of *p*. It is precisely in the act of believing that *p*, or coming to believe that *p*, that one *commits* oneself to the truth of *p*. Likewise, desiring *p* involves a certain mental commitment to the goodness of *p* that merely entertaining the goodness of *p* does not.

On reflection, the same kind of difference can be seen when we compare fearing a dog and entertaining the dog’s (fearsome-)dangerousness. The fear *commits* one to the dog’s dangerousness, whereas mere entertainment of the dog’s dangerousness does not. Sometimes contemplating a dangerous thing overmuch may *lead* us to suddenly experience fear regarding it. But the relation here is *causal* rather than constitutive: we move from one mental state (the contemplation) to a new, different mental state (the fear) – with the difference being precisely that only the new one embodies commitment to the thing’s (fearsome-)dangerousness. The proponent of attitudinal-representational properties has a natural explanation of this difference: fearing the dog involves representing-as-dangerous the dog, whereas entertaining the dog’s (fearsome-)dangerousness involves only representing the dog-as-(fearsomely-)dangerous.¹⁴



The representationalist may resist this line of thought by attempting to reductively account for the attitudinal-representational properties of fear and admiration in terms of those of belief and desire (see, e.g., Marks 1982, Gordon 1987). For example, she might suggest that fearing a dog is just a matter of believing that the

dog is present and dangerous and desiring that the dog be absent or undangerous. If this kind of reductive account could be carried out satisfactorily for all emotional attitudinal-representational properties, it would create a different kind of unity in our picture of mind, perhaps deeper than that which treats fear and admiration as on a par with belief and desire.

An immediate obstacle to this gambit is the apparent fact that the belief and desire appealed to in the account have *propositional* content, whereas the fear allegedly accounted for has non-propositional content. It is unclear how any combination of propositional attitudes can constitute a *non-propositional* attitude.

To overcome this immediate obstacle, the representationalist must adopt a more indirect reductive strategy: first account for fearing x in terms of fearing that p , then account for fearing that p in terms of the right combination of belief and desire. And likewise for all other apparently objectual, non-propositional emotional attitudes, such as love, admiration, and so on. However, both parts of this indirect strategy are quite dubious.

Consider first the notion that all emotions have propositional content. On this view, when, upon feeling a surge of love for my spouse, I say 'I love you,' my utterance reflects badly the structure of the conscious state it expresses. A less misleading utterance would take the form 'I love that _____.' But how are we to fill in the blank? 'I love that you exist' seems to express at most an aspect, or an implication, of what 'I love you' expresses. 'I love that you are lovely' is multiply problematic. For starters, as we saw it is far from clear what 'lovely' stands for here. More deeply, loving that a is F appears to imply taking a to be F , but a person in the clutches of a pathological relationship may not take her beloved to have whatever property 'lovely' is supposed to stand for. (By the same token, a person in the clutches of arachnophobia may fear a spider she does not take to be dangerous. It would then be plausible to say that she fears the spider but not that she fears *that* the spider is dangerous.) Furthermore, 'I love that you are lovely' seems to involve the kind of theoretical overkill by which we duplicate what makes a mental state one of love. The more accurate utterance would have to be 'I entertain that you are lovely.' But as noted, mere entertainment does not seem to involve that special emotional commitment that loving embodies. (I would be quite disappointed to hear that my spouse entertains my loveliness!) This is why we normally entertain the loveliness of so many more people than we actually love. It thus appears very hard to find a remotely plausible propositional rendering of 'I love you' – which tellingly should come as good news to all the beloved of the earth. Likewise for other objectual emotions (Forbes 2000, Montague 2007).¹⁵

So much, then, for the notion that all emotions have propositional content. Even if they did, however, the analysis of propositional emotions in terms of belief and desire is highly problematic. In particular, it faces a dilemma when it comes to the question of whether belief and desire are phenomenally conscious states. The mainstream view in the philosophy of mind of the past half-century has been that such states – certainly belief – have no proprietary phenomenal character. And while recent debates on so-called cognitive phenomenology (Bayne and Montague 2011) raise the specter of phenomenal consciousness in belief (and desire) anew, reductive representationalists have tended to insist that belief and desire are non-phenomenal states (Tye and Wright 2011). Plausibly, however, mental states such as fearing that the dog might bite you *are* (sometimes) phenomenally conscious – there is something it is like to experience such a fear.¹⁶ If believing that *p* and desiring that *q* have no phenomenal character, while fearing that *r* does have phenomenal character, it is hard to see how fearing that *r* could be *nothing but* the combination of believing that *p* and desiring that *q*. Even if just believing that *p* has no phenomenal character, while both desiring that *q* fearing that *r* do, it remains hard to see how fearing that *r* could simply consist in the combination of believing that *p* and desiring that *q*, given that fearing that the dog might bite you is so phenomenally *different* from desiring that the dog *not* bite you (and indeed from desiring that *q* for any *q*).¹⁷

For the strategy to have any chance, then, the representationalist must allow belief and desire to be phenomenally conscious states. A modest allowance grants belief a phenomenal character, but not a *proprietary* one. The phenomenal character of believing that *p* is exhausted, the story goes, by an auditory image of the sentence '*p*' recited in silent speech (or by a visual image of the sentence passing before the mind's eye) (see, e.g., Robinson 2006). This kind of phenomenal character seems insufficiently rich, however, to capture the feel of fear. It is one thing to claim that the phenomenology of believing that the dog is dangerous involves grasping the property of being dangerous, or grasping the concept DANGER, and that the phenomenology of grasping this property or concept is all there is to the phenomenology of fear. But here the claim is that the phenomenology of believing that the dog is dangerous only involves grasping the *sound* /'dānjər/ (or the shape d^a^n^g^e^r). This kind of purely auditory (or visual) phenomenology seems to have no resources for capturing the felt quality of fear. Subjectively, fearing something is not just 'hearing' sounds in one's mind's ear or 'seeing' shapes in one's mind's eye.

A more generous allowance is thus called for, granting belief and desire *proprietary* phenomenal characters. Once we do so, however, any pair of cognitive

and conative experiences, say consciously believing that there is whisky at the party and consciously desiring that there be whisky at the party, would already constitute an immediate counter-example to representationalism. For such a pair involves two phenomenally different states with the very same representational content. Moreover, the phenomenal difference between the two would appear to trace back to their distinctive attitudinal-representational properties. We have already noted that belief appears to involve mental commitment to the truth of what is believed, and desire a commitment to the goodness of what is desired, without truth and goodness showing up in their respective contents themselves.¹⁸ That is, belief is characterized by the attitudinal-representational property of representing-as-true, desire by that of representing-as-good. Given content identity, the phenomenal difference between conscious-believing that p and conscious-desiring that p seems to trace back to the difference between representing-as-true and representing-as-good.



In conclusion. Experiences such as fearing a dog or loving one's spouse appear irreducibly objectual, admitting of no reductive analysis in terms of propositional states such as belief and desire. Moreover, it is extremely odd, and incongruent with the rest of our picture of mind, to cast all such experiences as involving a single, uniform attitude of pure entertaining toward certain special properties. And finally, the whole idea that such emotional experiences are directed at properties, or states of affairs, or other entities wholly or partially non-concrete, is in any case incredible. Thus the hope of accounting for the identity conditions of all phenomenally conscious states entirely in terms of what these states represent runs into major obstacles at every turn. It is considerably more plausible to simply admit that phenomenally conscious states individuate partly in terms of *how* they represent, that is, in terms of distinctive attitudinal-representational properties they exhibit. Such attitudinal-representational properties not only exist, then, but are phenomenally manifest. So: among the phenomenal properties our conscious states instantiate are attitudinal-representational properties, properties that do not have to do with what our conscious states represent.

4. Conclusion: Accentuating the Positive

Our critique of representationalism is not merely destructive. It is also instructive. For it suggests a rather concrete picture of phenomenal consciousness. In this picture, phenomenal character is determined by the conspiracy of at least two factors: representational content and representational attitude or mode. What it is

like to fear a dog is fixed partly by the fact that a dog is what one fears and partly by the fact that fearing is what one feels toward the dog.

Once we recognize the role of attitudinal-representational properties in emotional experience, it becomes natural to posit them elsewhere in our theory of phenomenal consciousness. Compare a visual experience of a dog and an experience of *visualizing* a qualitatively indistinguishable dog (a dog with all the same visible surface features). Suppose that due to extraordinary circumstances, involving fogs and mirrors and much more besides, the degree of vivacity, clarity, precision, and resolution of the two experiences is identical. Hume might maintain that in that circumstance there is no phenomenal difference between the two experiences. This assumes that there is no phenomenal difference between vision and visualization *as such*. But it is not outlandish to think that these experiences do differ phenomenally, though not due to *what* they represent. Rather, they differ in virtue of *how* they represent. Perhaps the difference is something like this: the visual experience of the dog represents-*as-real* the dog, whereas the visualization of the dog represents-*as-unreal* (or perhaps represents-*as-imaginary*) the dog. Perhaps it is rather that the visual experience represents-*as-present* whereas the visualization represents-*as-absent* the dog. Perhaps the difference pertains to some other distinctive attitudinal-representational properties. The idea, in any case, is that the overall phenomenal character of both vision and visualization is fixed in part by attitudinal-representational properties these instantiate. I have not *argued* for this idea here, but the admission of a role for attitudinal-representational properties in phenomenal-state individuation in the emotional domain lends a minimal prima facie plausibility to the idea – the kind of minimal plausibility that recommends closer examination. And what is true of vision and visualization may be true, more generally, of perception and imagination.¹⁹

A similar contrast may attend perception and episodic memory. Consider two subjects, one of whom watching the rain falling and the other episodically remembering a qualitatively indistinguishable rain; suppose again the world conspires to equate the levels of vivacity, clarity, precision, and resolution in their respective experiences. Here too, there is a view worth examining that casts the two experiences as phenomenally type-different in virtue of their attitudinal-representational properties. The most natural version of this second view would be that while episodic remembering of the rain represents-*as-past* the rain, its perceiving represents-*as-present* the rain. Again, I do not *argue* for this view here, but do note its structural symmetry to the view I have argued for in the emotional domain.²⁰

Since POSITIVE is only an *existential* thesis, the case of emotional experience suffices to support POSITIVE. The case of imagination and episodic memory is a dialectical luxury I have not indulged here. At the same time, once we have admitted the ineluctability of attitudinal-representational properties in one area of phenomenal consciousness, there is not much of an additional cost in introducing them in other areas. Thus POSITIVE, supported by considerations of emotional experience, may inspire a unified theory of phenomenal consciousness that constitutes a substantial alternative to (reductive) representationalism.²¹

References

- Barwise, J. and J. Perry 1983. *Situations and Attitudes*. Cambridge MA: MIT Press.
- Bayne, T. and M. Montague 2011. *Cognitive Phenomenology*. Oxford: Oxford University Press.
- Block, N.J. 1990. 'Inverted Earth.' *Philosophical Perspectives* 4: 52-79.
- Block, N.J. 1996. 'Mental Paint and Mental Latex.' *Philosophical Issues* 7: 19-50.
- Byrne, A. 2001. 'Intentionalism Defended.' *Philosophical Review* 110: 199-240.
- Chalmers, D.J. 2004. 'The Representational Character of Experience.' In B. Leiter (ed.), *The Future for Philosophy*. Oxford: Oxford University Press.
- Crane, T. 2003. 'The Intentional Structure of Consciousness.' In A. Jokic and Q. Smith (eds.), *Consciousness: New philosophical perspectives*. Oxford and New York: Oxford University Press.
- Crane, T. 2009. 'Intentionalism.' In A. Beckermann and B. McLaughlin (eds.), *Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press.
- Dretske, F.I. 1981. *Knowledge and the Flow of Information*. Oxford: Blackwell.
- Dretske, F.I. 1995. *Naturalizing the Mind*. Cambridge MA: MIT Press.
- Dretske, F.I. 2000. 'Reply to Lopes.' *Philosophy and Phenomenological Research* 60: 455-459.
- Gordon, R. 1987. *The Structure of Emotion*. Cambridge: Cambridge University Press.
- Fine, K. 2001. 'The Question of Realism.' *Philosophers' Imprint* 1: 1-30.
- Forbes, G. 2000. 'Objectual Attitudes.' *Linguistics and Philosophy* 23: 141-183.
- Harman, G. 1990. 'The Intrinsic Quality of Experience.' *Philosophical Perspectives* 4: 31-52.
- Horgan, T. 1993. 'From Supervenience to Superdupervenience: Meeting the Demands of a Material World.' *Mind* 102: 555-586.
- Horgan, T. and J. Tienson 2002. 'The Intentionality of Phenomenology and the Phenomenology of Intentionality.' In D.J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford and New York: Oxford University Press.
- Klausen, S.H. 2008. 'The Phenomenology of Propositional Attitudes.' *Phenomenology and the Cognitive Sciences* 7: 445-462.
- Kriegel, U. 2002. 'PANIC Theory and the Prospects for a Representational Theory of Phenomenal Consciousness.' *Philosophical Psychology* 15: 55-64.
- Kriegel, U. 2015a. *The Varieties of Consciousness*. Oxford and New York: Oxford University Press.

- Kriegel, U. 2015b. 'Experiencing the Present.' *Analysis* 75: 407-413.
- Loar, B. 2003. 'Phenomenal Intentionality as the Basis for Mental Content.' In M. Hahn and B. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge MA: MIT Press.
- Lopes, D.M.M. 2000. 'What Is It Like to See with Your Ears? The Representational Theory of Mind.' *Philosophical and Phenomenological Research* 60: 439-453.
- Marks, J. 1982. 'A Theory of Emotion.' *Philosophical Studies* 42: 227-242.
- Montague, M. 2007. 'Against Propositionalism.' *Noûs* 41: 503-518.
- Pautz, A. 2010. 'Consciousness: A Simple Approach.' In R. Koons and G. Bealer (eds.), *The Waning of Materialism*. Oxford and New York: Oxford University Press.
- Robinson, S.W. 2006. 'Thoughts without Distinctive Non-Imagistic Phenomenology.' *Philosophy and Phenomenological Research* 70: 534-561.
- Shoemaker, S. 1994. 'Phenomenal Character.' *Noûs* 28: 21-38.
- Thau, M. 2002. *Consciousness and Cognition*. Oxford: Oxford University Press.
- Tye, M. 1995. *Ten Problems of Consciousness*. Cambridge MA: MIT Press.
- Tye, M. and B. Wright 2011. 'Is There a Phenomenology of Thought?' In Bayne and Montague 2011.

¹ Particularly influential in this has been an early paper by Harman (1990). For fuller developments of representationalism, see Shoemaker 1994, Dretske 1995, and Tye 1995. Other important defenses of the view are Byrne 2001 and Thau 2002.

² This has to do with a certain dissatisfaction with the notion of supervenience, as a philosophical tool, which has become widespread across many areas of philosophy. Crushing many subtleties, we may say that supervenience is not an *explanatory* relation, which makes it unsuited to carry, all by itself, the burden of various projects of philosophical explanation – much less the burden of *reductive* philosophical explanation. The fact that supervenience representationalism does not answer Q is but a symptom of that.

³ Some philosophers deny that (the right kind of) representational content can be fully understood in entirely non-phenomenal terms (Horgan and Tienson 2002, Loar 2003). When philosophers combine commitment to THIRD with this non-reductivism about representational content, they end up with non-reductive representationalism (Chalmers 2004, Pautz 2010).

⁴ This position comes in two varieties, depending on whether one holds (a) that both rectangularity and the-look-of-rectangularity (or the-sounds-of-rectangularity) show up in phenomenal consciousness or (b) that only the-look-of-rectangularity (or the-sound-of-rectangularity) does. This difference will not matter to our discussion.

⁵ A representationalist taking this line could also deny that there really is the audible-only feature of rectangularity, and on that ground deny that auditory shape experiences truly exist. This appears to be Dretske's (2000) response to Lopes' case. The idea is to concede that there are auditorily-based rectangularity beliefs, but deny that it follows there must be auditory shape experiences.

⁶ From the perspective of FEAR and ADMIRATION, the notion that danger and admirability appear in the *contents* of fear and admiration is an instance of what Barwise and Perry (1983) called 'the fallacy of misplaced information.'

⁷ Keep in mind that we are using representation talk in such a way that a hallucination of a lemon in a lemonless world does *not* represent nothing. In this usage, a mental state that represents nothing is not just a hallucination or a mental state for which there is nothing in the world that fits what it purports to be about; rather, it is a mental state that does not *purport* to be about anything.

⁸ Note that POSITIVE is a merely *existential* thesis, making no claim to the effect that *all* phenomenally conscious states owe their identity in part to the attitudinal-representational properties they exhibit. One reason for this measure of modesty is simply that, since THIRD is a universal claim, an existential counter-claim is enough to undermine it.

⁹ The two difficulties are independent of each other. The claim that no entity lends itself only to conscious representation applies *within* each mode of experience. It states that it should be possible, for any property which is represented in an auditory experience, to imagine that very property being represented in 'subliminal' auditory perception; that any property which can be represented by a conscious flair of anger should be also available for representation in suppressed anger as well; and so on. Here the specific mode plays no special role. Indeed, the problem would apply even if there were no experiential modes. In contrast, the difficulty to be developed in the argument from non-perceptual modes concerns the relations between different modes of experience.

¹⁰ Consider: (FOURTH) For every phenomenally conscious state S, there is a class of representational properties R, such that what makes S the phenomenally conscious state it is, and a phenomenally conscious state at all, is that S instantiates the members of R it does, and instantiates members of R at all. (Unlike THIRD, FOURTH is *consistent* with POSITIVE, insofar as attitudinal-representational properties are representational properties.)

¹¹ The distinction is drawn before Chalmers for the question of perceptual experience by Block (1996), who calls impure representationalism 'quasi-representationism,' and Byrne (1997), who calls it 'intramodal intentionalism.'

¹² I am assuming here that there is a close link between reductivism and purism. Recall that according to REDUCTIVE, the representational content of phenomenally conscious states can be accounted for in terms of non-phenomenal facts, that is, facts no constituent of which is phenomenal. The main representationalist strategy in this area attempts to account for representational content in terms of causal, informational, and/or teleological relations between conscious states and worldly entities. Even granting the success of this reductive enterprise, impure representationalism would insist that a reductive philosophical explanation of phenomenal consciousness would also require a reductive account of attitudinal-representational properties. Of course, none of this *rules out* the existence of some *other* reductive strategy, more suitable for attitudinal-representational properties. But the point I hope to establish here is only that the *representationalist* reductive strategy – appealing centrally to informational semantics, teleosemantics, and so on – runs into a dead end once we move from pure to impure representationalism.

¹³ I underline the occurrence of 'possible' in this statement – there is no claim of infallibility here!

¹⁴ The representationalist, in contrast, must deny that fearing a dog is different from entertaining the dog's (fearsome-)dangerousness. In doing so, she appears to introduce a major disunity into our picture of the mind. If belief and desire, as *the* paradigmatic mental states, have distinctive attitudinal-representational properties, then a picture where fear and admiration have their own is *mutatis mutandis* more unified than a picture where they do not.

¹⁵ In the domain of perception, too, we tend to distinguish perceiving *x* from perceiving *that p*. The notion that we can somehow reduce the former to the latter appears very much a non-starter. Crucially, perceiving *x* does not require the possession of any concepts, whereas perceiving *that p* does (Dretske 1993). An infant who encounters a squirrel for the first time sees *the squirrel* but does

not see *that it is a squirrel*, lacking as she does the concept of a squirrel. If the infant is old enough, she may perceive *that* there is a reddish-brown, squirrely-shaped thing before her. But the infant may *not* be old enough – she may be so young that she lacks even a concept for the relevant color (perhaps the squirrel dancing before her is the first object she encounters which has any red or brown in it), not to mention a concept for the relevant shape. Thus it is highly unlikely that perceiving *x* could be satisfactorily analyzed in terms of perceiving that *p*. It would be surprising, then, if loving *x* could be analyzed in terms of loving that *p*, admiring *x* in terms of admiring that *p*, fearing *x* in terms of fearing that *p*, and so on.

¹⁶ Compare: there is something it is like to see *that* the squirrel is reddish-brown (Clausen 2008).

¹⁷ Note that in the present dialectical context the representationalist *cannot* respond that phenomenal character of fear that *p* simply derives from, or consists in, the phenomenal character of some accompanying fear of *x* (e.g., that the phenomenal character of fearing that dog will bite you consists in that of fearing the dog or the dog's potential bite); for the representationalist strategy under consideration attempts precisely to account *without remainder* for such apparently objectual, non-propositional fear in terms of belief and desire. The strategy is *doomed* if the representationalist concedes that fearing *x* has a phenomenal character whereas belief that *p* does not.

¹⁸ It would indeed be odd to place truth/obtaining and goodness in the *contents* of belief and desire. In believing that there is whiskey at the party, *what* one believes is simply *that there is whisky at the party* – not that *it is true* that there is whisky at the party (nor that *there being whisky at the party obtains*). Likewise, in desiring that there be whisky at the party, *what* one desires is *that there be whisky at the party* – not that *it be good* if there be whisky at the party. Thus belief's truth-commitment and desire's goodness-commitment do not appear to come from *what* belief and desire represent, but rather from *how* they represent – from the distinctive ways these states have of framing their contents.

¹⁹ I argue for this line on the perception/imagination distinction in Kriegel 2015a Ch.6.

²⁰ I argue for this line of thought more fully in Kriegel 2015b.

²¹ Work on this paper was supported by the French National Research Agency's grants ANR-11-0001-02 PSL* and ANR-10-LABX-0087, as well as by grant 675415 of the European Union's Horizon 2020 Research and Innovation program. For comments on a previous draft, I am grateful to David Chalmers.