Philosophy and
Phenomenological Research

Philosophy and Phenomenological Research
doi: 10.1111/phpr.12395

Valence and Value

PETER CARRUTHERS
University of Maryland

Valence is a central component of all affective states, including pains, pleasures, emotions, moods, and feelings of desire or repulsion.[1]This paper has two main goals. One is to suggest that enough is now known about the causes, consequences, and properties of valence to indicate that it forms a unitary natural-psychological kind, one that seemingly plays a fundamental role in motivating all kinds of intentional action. If this turns out to be true, then the correct characterization of the nature of valence becomes an urgent philosophical issue. There appear to be just two accounts that have the required generality. According to one, valence is a nonconceptual representation of value. According to the other, valence is an intrinsic qualitative property of experience. (Both views maintain that valence is directly motivating.) The second goal of the paper is to contrast and evaluate these two views of the nature of valence, drawing on the relevant empirical findings. Overall, I suggest that the representational account is more plausible.

1. Affect, Valence, and Motivation

Cognitive science regards affective states as a broad class, grouping together phenomena that common sense treats as belonging to very different kinds. The class includes: headaches and orgasms; feelings of longing and repulsion; emotions of anger, fear, disgust, amusement, and grief; feelings of enjoyment or boredom in an activity; and moods such as happiness, sadness, and depression. Some affective states, such as fear and anger, are propositional attitudes. (One is afraid *that* the bear will attack, or angry *that* one's colleague has made a cutting remark.) But others are not. (Neither an orgasm nor a depressed mood appear to be *about* anything.) It is widely agreed, however, that all affective states share two dimensions of *valence* and *arousal* (Russell, 1980, 2003; Reisenzein, 1994; Rolls, 1999). All affective states have either positive or negative valence (positive for orgasm, negative for fear); and all can be placed along a continuum of bodily arousal (high or low heart-rate, speed of breathing, tensing of muscles, and so on).

There has been some debate among cognitive scientists over how affective states are individuated. (This is also a question that has exercised philosophers, especially in the domain of emotion; Solomon, 1984; Griffiths, 1997; Prinz, 2004). Some have claimed that the varying positions occupied by affective states on the valence and arousal dimensions are all that really distinguish them (Russell, 2003). This is the so-called "core affect" view. Others have emphasized that affective states are characterized by distinctive

[1] Positive and negative valence are sometimes described as "pleasure" and "displeasure" respectively. I use the term "valence" throughout, both for theoretical neutrality and to emphasize that our discussion is not about vernacular concepts.

forms of appraisal of environmental or bodily events, such as an appraisal of something as *threatening* in the case of fear (Ellsworth & Scherer, 2003). And yet others have stressed that different bodily postures, facial expressions, and action-tendencies are also characteristic of different kinds of affective state (LeDoux, 2012).

Whether or not affective states can be *individuated* (in part) via their distinctive action-tendencies and patterns of appraisal, most affective episodes include both. All emotions and desires, at any rate, automatically activate motor plans (for approach, in the case of desire, for swift retreat in the case of fear) that need to be inhibited by executive signals if those plans are not to be acted on. And all affective states result from assessments of the relevance of environmental or bodily events, either to previously formed goals, or to one's underlying values stored subcortically as dispositional properties of reward-systems in the basal ganglia. (The values in question, here and throughout, are understood to be subjective ones, of course.) These appraisals are generally swift and unconscious, operating at many different levels of processing of the sensory input. (For example, they attract attention to evaluatively-relevant but currently-unconscious environmental stimuli; Corbetta et al., 2008.) Yet they can also involve *re*appraisals of a stimulus, either by looking closer, for example (after attention has been drawn to it), or through strategically deployed re-representings of the stimulus, of the sort that are involved in top–down strategies for emotional self-management (Gross, 2015).

My question is not about individuation. It is rather about what all affective states have in common—and more specifically, the valence dimension they all share. It seems increasingly likely that valence constitutes a single natural-psychological kind, the same in nature across all the different varieties of affective state. Valence-processing appears to be underlain by a single (albeit multicomponent) neurobiological network, involving not just subcortical evaluative regions in the basal ganglia, but also the anterior insula and anterior cingulate, together especially with orbitofrontal and ventromedial prefrontal cortex (Leknes & Tracey, 2008; FitzGerald et al., 2009; Plassmann et al., 2010; Bartra et al., 2013). The latter regions are the primary projection areas for valence signals in the cortex. These signals are thought to provide an evaluative "common currency" for use in affectively-based decision making (Levy & Glimcher, 2012). Valence produced by many different properties of a thing or event can be summed and subtracted to produce an overall evaluative response, and such responses can be compared to enable us to choose among options that would otherwise appear incommensurable.[2]

Moreover, not only can grief and other forms of social suffering be blunted by using Tylenol, just as can physical pain (Lieberman & Eisenberger, 2009; Lieberman, 2013), but so, too, is pleasure blunted by the same drugs (Durso et al., 2015). In addition, both pain and pleasure are subject to top–down placebo and nocebo effects that seemingly utilize the same set of mechanisms. Just as expecting a pain to be intense (or not) can influence one's experience accordingly, so can expectations of pleasure increase or decrease the extent of one's enjoyment (Wager, 2005; Plassmann et al., 2008; Ellingsen et al., 2013). Indeed, moderate pain that is lesser than expected can even be experienced as

[2] Although it is not my focus in this paper, valence also serves as a teaching-signal for evaluative learning (Schroeder, 2004). If one experiences something to be more pleasant than expected, the stored values associated with that thing are ratcheted upwards a notch. Likewise, if it is less pleasant than predicted, stored values are adjusted downwards.

pleasant, suggesting the involvement of a single underlying mechanism (Leknes et al., 2013).

If valence constitutes a single natural kind, the same across all different types of affective state, then this places strong constraints on accounts of specific affective states like pain. An adequate account of the hurtfulness of pain should bring out what it has in common with the valence of any other negative affective state, such as fear or grief. This has not been adequately appreciated in the philosophical literature, where it is common to analyze the hurtfulness component of pain in terms that could not apply to all negative affective states. For example, Cutter & Tye (2011) analyze it in terms of *harmfulness*, which fails to apply to sadness or depression. The present paper will take for granted that valence is a unitary kind, and will evaluate competing accounts of the nature of that kind. But before getting to that, more needs to be said about the role of valence in motivating action. For if, as I shall suggest, all intentional action is psychologically dependent on valence, then characterizing the unitary nature of valence becomes quite urgent.

It is widely believed by affective scientists that valence is intrinsically motivating, and plays a fundamental role in affectively-based decision making (Gilbert & Wilson, 2005; Levy & Glimcher, 2012). When we engage in prospection, imagining the alternatives open to us, it is valence-signals that ultimately determine choice, generated by our evaluative systems responding to representations of those alternatives. The common currency provided by these signals enables us to compare across otherwise incommensurable alternatives and combine together the values of the different attributes involved. Indeed, there is some reason to think that valence might provide the motivational component underlying all intentional action, either directly or indirectly. Or so I shall now briefly argue.[3]

It might appear that the so-called "somatic marker hypothesis" championed by Damasio (1994) and others conflicts with the claimed foundational role for valence in decision making—suggesting, on the contrary, that it is the arousal component of affect that plays the primary role. But in fact it is likely that Damasio merely endorses a hedonic, self-focused, construal of valence (of the sort to be discussed later in this article), as do many of the psychologists who work on affect. At any rate Damasio, too, emphasizes the crucial role of orbitofrontal cortex and ventromedial prefrontal cortex in the motivational component of human decision making, and these are widely thought to be among the primary projection areas for valence signals in the brain (as opposed to arousal ones, which are represented in somatosensory cortex and elsewhere) (Levy & Glimcher, 2012).

Although everyone in the field of affective science will agree that valence is *important* for motivating intentional action, it is less clear that it is *essential*.[4] Two issues are worth

---

[3] It isn't true that *all* forms of action are motivated by valence. For some, like habitual actions, can be triggered and controlled by perceptual states in the absence of motivation (Lisman & Sternberg, 2013). Moreover, the action-tendencies, facial expressions, and bodily postures that are characteristic of many emotions and moods seem to be caused directly by subcortical evaluative systems, independently of valence-based decision-making. Indeed, I suggest that the distinction drawn in the empirical literature between *liking* and *wanting* (Berridge & Kringelbach, 2008) is really a difference between pleasure at the thought of doing something, which may lead to a decision to do it (liking), and primitively-caused approach or consummatory behavior (wanting). But this leaves open as a possibility that all *intentional* actions (actions that are caused by decision-making processes) are grounded in the common currency of valence.

[4] Note that the necessity in question here is psychological, not conceptual or metaphysical. The question is whether all intentional—that is, decided-upon—actions are *causally* dependent upon valence *in beings like us*.

brief discussion. One is whether goals and intentions can motivate action independently of affective states. The other is whether cognitive states alone (for example, *beliefs* about what is good or bad) can motivate action. Let us take these questions in turn.

It is an important part of the functional role of intentions that they can initiate and control behavior in the absence of affective states or affective processing (Bratman, 1987, 1999). Moreover, intentions can constrain and foreclose affect-involving practical reasoning. Likewise, one's goals can issue in behavior without requiring support from one's affective states. Notably, both intentions and goals form parts of the brain's control network, located especially in dorsolateral prefrontal cortex (Seeley et al., 2007). Note that this network is distinct from—although often interacting with, of course—the affective networks located in ventromedial prefrontal cortex and subcortically in the basal ganglia.

Although goals and intentions can motivate action independently of current affect, it may be that such states nevertheless depend on affect at the stage when they are initially formed. Many in the field think that one adopts a goal, or forms an intention, by reflecting on and responding affectively to the available options (Gilbert & Wilson, 2005, 2007). In fact, it may be helpful to think of goals and intentions as existing on a spectrum of more-or-less abstract *motor* plans. One can have the intention of grasping a cup to drink, one can intend to attend a friend's wedding, or one can have the goal of becoming a property owner. In each case one can select among a range of potential implementations, prompted by affordances in the environment and other factors. And in each case the mental state in question will have been formed from previous affectively-laden decision-making processes, while thereafter being capable of initiating action in the absence of such processes. So the actions one performs are still dependent on valence in a distal if not in a proximal sense.

We turn now to the question whether evaluative *beliefs* can motivate action in their own right. There is significant evidence that they cannot. In particular, people with damage to orbitofrontal and ventromedial prefrontal cortex, which are the main targets for valence signals in the cortex, can make perfectly sensible *judgments* about what it would be good to do; but their actual decision-making goes all to pieces (Damasio, 1994; Bechara et al., 2000). Likewise, people with flattened affect can know, intellectually, what it would be good to do while feeling no corresponding impulse to do it. However, this is not to say that value beliefs cannot influence motivation indirectly, via their downstream influence on affective processing. In fact, there are a number of ways in which this can happen.

Suppose that moral testimony leads one to acquire a belief that it is good to be kind to insects. Initially one might feel little or no inclination to do what one believes to be good. But knowing that people who do what others believe to be good are admired and supported by their communities, and responding affectively to the thought of the latter outcome, one might come to have a positive affective reaction to the thought of being kind to insects. This may be the standard route through which one internalizes the values of one's community (Sripada, 2007).

In contrast, consider a moral rebel, who reasons his way to a belief about what is good that is wildly at odds with the beliefs of the community. (Peter Singer might provide a real-life example.) Is there any way for such a belief to cause the activity in question (not eating meat, as it might be) to become positively valenced, and hence issue in motivation? There are a number of possible routes through which this might happen, in

fact. For example, one might appraise acting on one's belief as likely to make one famous or prestigious (albeit unconsciously, no doubt), leading the positive valence directed at the latter to become transferred to the actions themselves. Or one might admire (responding with positive valence toward) those who stand by their ethical principles. Then appraising the action of no longer eating meat in such terms may cause it to become positively valenced.

More simply, however, beliefs about what is good can give rise to affective responses directly. This is because of the widespread phenomenon of predictive coding (Clark, 2013), which in this case leads to an influence of top–down expectations on affective experience. We know that expecting an image to depict a house can make it appear more house-like than it otherwise would (Panichello et al., 2013). And likewise, expecting something to be good can lead one to experience it as more valuable than one otherwise would. This is the source of placebo-effects on affective experience (Wager, 2005; Plassmann et al., 2008; Ellingsen et al., 2013). Just as expecting a stimulus to be a house can cause one to experience it as house-like even if it is, in fact, completely neutral or ambiguous, so believing something to be good may lead one to experience it as good in the absence of any initial positive valence.

Similarly, the mere belief that one has chosen one thing over another—thereby acquiring an implicit belief that one *prefers* the one to the other—will generally lead to heightened positive affect directed at the item one believes oneself to have chosen (Harmon-Jones & Mills, 1999; Lieberman et al., 2001). Indeed, this remains true even if one never really made a choice, but has been duped by cunning experimenters into believing that one did (Sharot et al., 2008; Johansson et al., 2014). Moreover, the effects of believed-choice on affect are still detectable three years later (Sharot et al., 2012).[5]

It may be, then, that the valence component of affect plays a fundamental and psychologically-essential role in motivating intentional action. It is the ultimate source of the decisions that issue in intentions for the future and the adoption of novel goals. And it is through the effects of evaluative beliefs on valence-generating value systems that the former can acquire a derivative motivational role. If these claims are correct, then understanding the nature of valence is crucial for understanding both decision-making and action.

In what follows, therefore, I shall make two assumptions about the nature of valence. (I don't pretend to have defended either of these assumptions sufficiently here.) One is that valence is a natural kind, the same across all different forms of affective state. The second is that valence is directly motivating. Positive valence motivates one to pursue the valenced object or event; negative valence motivates one to reject it. There appear to be just two kinds of account that have the degree of generality required to substantiate these assumptions. Section 2 will introduce and begin to compare them.

2. The Nature of Valence: Two Views

We can begin by considering the nature of pain, which will form one of our central examples of affective experience throughout. As is now widely known, pains have both a sensory and a valuational component, mediated by distinct neural pathways. People in

[5] Note that the influence of belief, here, is not just behavioral. Rather, changes can be detected deep within affective value-processing networks in the brain. See any of the fMRI studies cited in this and the previous paragraph.

pain respond to certain kinds of morphine by saying that their pain *sensations* are unchanged—feeling just as sharp, say, and located in the same position in the body—but that they no longer *care*. Their pains are no longer *hurtful*. (Likewise for people who have undergone anterior cingulotomy as a treatment for chronic pain; Wilkinson et al., 1999.) The valuational aspect of pain can be identified with negative valence. But what of the sensory component? Here I assume a representational account. Pain sensations are nonconceptually-represented properties of one's body. Pain sensations have a represented location (in one's foot, say, or in one's elbow), while also having represented intensities and other spatial and temporal features (stabbing, say, or throbbing).[6] The experience of pain thus generally involves both a sensation of pain together with negative valence. Moreover, each of the two views of valence to be considered here agrees that the latter is in some sense consciously *felt*, as we will see.

On one view (the view I advocate) the valence component of affective states like pain and pleasure is a nonconceptual representation of badness or goodness. The valence of pain is a fine-grained perception-like representation of *seeming badness* and the valence of pleasure is a similarly fine-grained representation of *seeming goodness*, where both exist on a single continuum of *seeming value*. However, these phrases need to be understood in a way that does not presuppose any embedding within the experience of the concepts BAD and GOOD. One has to *use* these concepts in describing the content of a state of valence, of course (just as one has to use color-concepts in describing the content of color experience), but that doesn't mean that the state in question either embeds or presupposes the relevant concept.

For comparison, consider nonconceptual representations of approximate numerosity, of the sort entertained by infants and nonhuman animals (and also by adult humans) (Barth et al., 2003; Jordan et al., 2008; Izard et al., 2009). In describing the content of such a representation one might say something like: the animal sees that there are *about thirty* dots on the screen. This needs to be understood in a way that carries no commitment to the animal possessing the concept THIRTY, however. Rather, what we now know is that the representation is more like a continuous curve centered roughly on thirty that allows the animal to discriminate thirty dots from forty dots, for example, but not thirty from thirty-five.

On one view, then, the valence component of pain consists in the nonconceptually-represented *seeming badness* of a concurrent sensation of pain. When one's foot hurts, one experiences a particular sensory quality in one's foot (intense and stabbing, say) and one experiences that sensory quality as (nonconceptually) bad. Generalizing, one can say that the valence component of any affective experience is a fine-grained, nonconceptual, representation of the goodness or badness of the object of that experience. (I shall refer to this as "the representational account" of valence.) These representations are amodal ones (that is, nonsensory, or not specific to any perceptual modality), and are therefore not properly described as forms of sensory experience themselves. Nevertheless, they seem to make a distinctive difference to one's ongoing conscious experience. There is something that it is like to experience agonizing pain, and this is very different from what it is like to experience the same pain sensations after one has undergone anterior

[6] For present purposes it can be left open whether the sensory component of pain is best thought of as a secondary quality represented at some specific location in one's body (as I am inclined to think), or whether the content represented is a primary-quality physical disturbance of some sort (as Tye, 2006, claims).

cingulotomy, when one is no longer bothered by them. Moreover, on the view I propose, valence representations are intrinsically motivating. Positive valence motivates one to obtain the object that seems good to one, whereas negative valence motivates one to avoid or get rid of what seems bad to one.

What is the relationship between nonconceptual representations of value, embodied in one's affective states, and conceptual ones, located in one's beliefs? What is the relationship between *seeming good* and *believing good*? I suggest that positive valence directed towards something causes at least a temptation to make the corresponding judgment. Just as seeing leads to believing by default, so does seeming-good lead to believing-good by default. Just as a nonconceptual representation of a red surface will generally lead to belief that the surface is red (unless inhibited by top–down considerations), so a positively-valenced reaction that leads something to seem good to one (nonconceptually) will normally issue in a belief that it *is* good.[7]

Why do I propose that valence is a *non*conceptual representation of value, rather than a conceptual one? Why not identify positive valence with the belief that something is good, and negative valence with the belief that something is bad? One source of support for the proposal is that it enables us to explain how valence can be phenomenally conscious, on the assumption that only states with nonconceptual content are phenomenally conscious. I will return to this idea in Section 4. But another motivation is that valence can then be what *grounds* our judgments of *better* and *worse* when we engage in prospective reasoning and decision making. (This is the "common currency" role of valence again.) Just as nonconceptual gradations in redness can ground a judgment that one object is redder than another (although both are red), so the valence associated with two good outcomes can ground a judgment that one is better than the other.

On the view I favor, then, valence is a nonconceptual representation of value. The main competitor for this account holds that the valence component of pain (and of other affective states) is a distinct qualitative property that pain *experiences* possess, where this property is, somehow, regarded as intrinsically bad. (I shall refer to this as "the hedonic account" of valence.) Alongside (and normally attaching to) the experienced sensory aspect of pain there is an intrinsic (non-representational, non-relational) quality that is intrinsically bad or unwelcome, and which thus motivates one to do things to remove or ameliorate the experience in question. Likewise with positive valence: when one contemplates eating a piece of chocolate cake after dinner one experiences (in addition to the thought of eating the cake together with the visual and gustatory images this entails) a qualitative component of one's experience that is felt to be intrinsically good. This is taken as a signal that eating some cake would produce further experiences of just that sort, thus motivating one to go to the fridge for a slice. In effect, the idea is that there are intrinsic value *qualia* that inhere in our affectively-laden experiences.

One initial strike against the hedonic account is that accords less well with our affective phenomenology than does the representational one. When a bear looms out of the bushes while one is hiking (causing fear) it is the threatening aspect *of the bear* (its size, its claws) that seems bad. All of one's focus when afraid is generally outward-directed,

[7] Some philosophers deny that there is such a thing as nonconceptual content, of course (e.g. McDowell, 1984). But they nevertheless accept that there is a distinction between what they take to be the fine-grained indexical concepts employed in experience and the coarser-grained concepts employed in thought. Everything I say here could be cast in such terms if necessary.

targeted on the object of one's fear. And that is what (according to the representational account) seems bad as a result. Negative valence represents the presence of the bear (or its threatening aspect) as bad. Similarly, when a colleague says something insulting at a meeting and one feels anger: one's focus is outward, on the person's action, and it is the insult itself that seems bad. Moreover, desires as well as emotions are generally outward focused: it is the chocolate cake itself (or perhaps better, the act of eating it) that seems good when one gazes longingly at it; and it is fame itself that seems good when one imagines being (and wants to be) famous.

In contrast, the hedonic account gives a much less natural, more self-focused, treatment of these cases. The sight of the bear causes an intrinsically-bad quality to become a component of one's experience, and the presence of the bear is only taken to be bad because one believes that it is causing this quality (while also believing that the bear may do things to cause other experiences containing qualities of the same sort, of course). Similarly, the colleague's insult causes an unpleasant quality to enter one's experience (thereby *making one feel* bad), and the insult itself is only derivatively bad, because one believes that it caused the presence of a bad experience.

On the other hand, it might seem that the hedonic account is especially well positioned to handle the valence component of moods like sadness or depression. When one is depressed, it might be said, one just *feels bad*, without there being any object or event that one feels bad *about*, or *represents as* bad. But in fact even moods can have a basic outward focus. It is *the world* that generally seems flat, colorless, and empty of meaning to those suffering from depression (Lambie & Marcel, 2002). This is easily handled by an account of negative valence according to which whatever one attends to is nonconceptually represented as bad. One merely has to note that depression is a state of persisting negative affect, leading all objects of attention to seem equally bad and uninviting. But even if we suppose that some forms of depression are *not* worldly-focused, this needn't lend support to a hedonic account of valence. For it can be one's bodily state—one's lassitude, one's slumped posture, one's inability to summon the energy to do anything—that is experienced as (that is, nonconceptually represented as) bad.

In any case it is obvious that not all valenced states need have a *worldly* focus. Pains and orgasms lead one to focus on the relevant sensory qualities of one's body. On the representational account, it is the sensation of pain that seems bad, and the sensations present in one's genitals during orgasm that seem good, respectively. Even in these cases, however, it is one's *sensations* that are evaluated as good or bad, not (in the first instance) one's *experience of* those sensations. When one has acute pain in a gouty toe, for example, it is the sensation in one's toe that feels bad. "Make *that* go away!" one might say, gesturing toward the toe (meaning not the toe itself, of course, but the sensations felt as present within it).

The hedonic account of the hurtfulness of pain, in contrast, is meta-experiential. It is one's *experience of* a pain sensation that is supposed to be intrinsically bad (because containing a bad intrinsic property or negative *quale*). The *sensation* of pain (the represented property of one's body) is not bad, except insofar as the *representing* of it issues in a bad experience. Nor is the threatening bear itself bad, except insofar as the experience of fear that it causes is bad. One problem for such a view, then, is that it seems to require considerable sophistication on the part of an affective experiencer, including a capacity to be aware of one's experiences as such. For how could one be motivated to get rid of an intrinsic property of one's experience (the valence component of pain) unless one were

aware of it *as* an experience (and hence were capable of representing it). Yet a great many animals (as well as human infants) can feel pain and fear, one might think; and many seem capable of valence-based prospective planning (Bird & Emory, 2009a, 2009b; Taylor et al., 2010; Hanus et al., 2011; Völter & Call, 2014). This is easy to understand if the valence component of affect represents a state of the body or environment as nonconceptually good or bad. It is harder to swallow that all creatures capable of affect and/or prospective reasoning are also meta-aware of their own experiences as such, in which bad qualia inhere.

A hedonic theorist can push back against these criticisms by asking, "Is it the chocolate cake itself that is good, in the first instance, or one's experience of the chocolate cake?" (Likewise for pains and orgasms.) For there is a natural tendency to respond, "One's experience of the cake!" (Aydede, 2014). What is surely true is that the cake is no good to one when *un*experienced. And likewise there is nothing bad about the bear in the bushes if one fails to see it, and passes by uneventfully without having been afraid. But this does nothing to support the hedonic account. Since valence is a nonconceptual representation produced by online evaluative processing, it only exists when it is a component of an ongoing experience of some sort. While the badness present in an experience of pain or of a bear, and the goodness present in orgasm or in eating chocolate cake, are experience-dependent, what is represented *as* bad or good in each case needn't be one's experience as such, but the object of one's experience.

Of the two views of valence just sketched, then, there is some reason to prefer the representational account. Further considerations in support of this approach will emerge in what follows, as we consider the implications of the two views for debates about hedonism (in Section 3) and the nature of phenomenal consciousness (in Section 4).

3. Valence and Hedonism

As we noted in Section 1, cognitive scientists have increasingly come to recognize the vital role played by *prospection* in human practical reasoning and decision making (Damasio, 1994; Gilbert & Wilson, 2005, 2007; Buckner, 2010; Seligman et al., 2013). When contemplating a decision or choosing among options one envisages (that is to say: one imagines) the actions and outcomes in question, responding affectively to them. The contents in question are appraised for relevance and one's value systems respond with positive or negative valence directed at the alternatives in question. On the account of valence proposed here, the result is that those options seem good, or seem bad, or seem better or worse, leading one of them to be selected. Note that practical reasoning, on this account, is almost entirely world-focused. It is the options and outcomes being contemplated that are represented as nonconceptually good or bad, leading them to seem attractive or unattractive, and thereby providing intrinsic motivation to pursue or avoid them.

The hedonic account of valence provides a very different picture of human practical reasoning. Indeed, when combined with the assumption defended briefly in Section 1—that all decision-making is ultimately affect-based—it leads straight to a form of motivational hedonism.[8] When one envisages a future choice or outcome, the representations in question are

[8] The arguments presented in this section can remain cogent even if not all decision-making is affect-based. But they would need to be re-cast slightly. It would just need to be claimed (as is surely true) that the sorts of examples I consider (such as sacrifice for a cause, or to save the life of a loved one) *can* involve valence-based decision-making.

processed by one's evaluative mechanisms, and some degree of positive or negative valence results. This alters the intrinsic character of one's experience while one envisages the future events, leading *the experience* to seem intrinsically good or bad. One then treats these feelings as a signal relating to the future event, believing that if it were to occur one would experience a similarly good or bad feeling (although one that is perhaps more intense and long-lasting). This leads one to select or avoid the act in question. Ultimately, then, all actions are undertaken in order to provide oneself with good or bad experiences.

This hedonistic construal of prospection is widespread in the empirical literature. Here, for example, is a quotation from Gilbert & Wilson (2007):

> People mentally simulate future events, but how do they use those simulations to predict the event's hedonic consequences? As the mere thought of eating a liver popsicle reveals, mental simulations of the future can elicit hedonic reactions in the present. People use their immediate hedonic reactions to simulations as predictors of the hedonic reactions they are likely to have when the events they are simulating actually come about. People do not imagine feeling anxious while having a colonoscopy so much as they imagine a colonoscopy, feel anxious, and then take this anxiety as an indicator of the feelings they can expect to experience during the procedure itself. Simulations allow people to "preview" events and to "prefeel" the pleasures and pains those events will produce. (1352; references omitted.)

The representational account would describe these cases quite differently. In particular, when one imagines the colonoscopy and feels anxious, the negative valence thereby produced *represents* the *colonoscopy* as (nonconceptually) bad. It may be, of course, that the reason why one is anxious at the thought of undergoing a colonoscopy is that one expects it to be painful, and it is the negative-valence component of anticipated pain that makes the colonoscopy seem bad. Still, it is one thing to say that one opts not to have a colonoscopy because *one expects it to be painful and the pain is represented as bad* (this is the representational account), and it is quite another thing to say that one avoids the colonoscopy because the thought of it *makes one now feel bad*, and one predicts that the colonoscopy would produce similar experiences (this is the hedonic qualitative-character account).

There are notorious difficulties with motivational hedonism, of course (Sober & Wilson, 1999), although it is also notoriously difficult to refute definitively (Stich et al., 2010). Most obviously, perhaps, people sometimes act to achieve outcomes they know they will never see realized. The most dramatic cases are people who sacrifice their lives for a cause. Seen from the perspective of the representational account of valence such examples are easy enough to understand. One imagines the state of the world once the revolution has been achieved (or whatever) and it seems very good, more good than does one's own continued existence in the absence of the revolution. The sacrifice of one's life is therefore comprehensible. But from a hedonist perspective the case becomes much more puzzling. For no matter how good one feels when one imagines that the revolution has arrived, one cannot rationally take this as a signal that one will feel comparably good when it *does* arrive if achieving the revolution would cost one one's life. (At least, this is so provided one doesn't both believe in an after-life and believe that one will feel especially good in the after-life knowing that one did one's bit for the revolution.)

At this point a hedonist can opt for either of two alternative replies. The first would be to claim that the mechanisms that produce positive affect at the thought of the revolution *also* produce a belief that is resistant to top–down influences (and

is presumably unconscious), to the effect that one *will* feel good when the revolution arrives. The motivation to sacrifice one's life therefore persists even when one acknowledges at a conscious level that one will experience nothing thereafter. This is possible, but looks ad hoc, and lacks any independent motivation. Indeed, it seems inconsistent with the known permeability of affective systems to top–down influence.

The other option would be to focus rather on the affective state one predicts for oneself if one does *not* sacrifice oneself for the revolution. Perhaps when one envisages such a future one anticipates feeling so guilty and miserable that death would be preferable. But this, too, would depend on a false belief. For we know that people generally adjust quite quickly to the choices they have made, rationalizing those choices to themselves after the fact (Moller, 2011; Wilson, 2011). Nor does it seem likely that the would-be revolutionary's motivation depends on any such belief. Even if convinced of the reality of emotional resilience one can imagine such a person responding, "But it isn't about *me* at all; it is about achieving justice for everyone!"

Moreover, given the widespread (and surprising) extent of resilience in the face of loss (Moller, 2011), it looks like a hedonist is committed to claiming that many of the choices that people make are actually prospective reasoning errors. Consider someone who takes out an expensive second mortgage on his home, or who donates a kidney, to save the life of a loved one. The hedonist's construal of the reasoning involved is that life without the loved one is predicted be much less pleasurable than life with little money. But given that the person would probably adapt quite swiftly to the loss of the loved one, this may be a mistake. In fact, he may be happier keeping the money (or kidney) and losing a spouse. From the perspective of the representational account, however, there is no error. The loss of the loved one *seems very bad*, much worse than a life with little money; and this may appropriately reflect one's underlying values. It is because one *values* the loved one that one makes the sacrifice, not because one wants to feel good (or to feel less bad).

There is also a more general theoretical difficulty with the hedonic account of practical reasoning. This is that there is a stark mismatch in *type* between the contents of the appraisals that issue in positive or negative valence (which are generally world-directed) and the content of one's resulting experience. But to set up this point we first need to draw a distinction between input-content and output-content, which both theories of valence should recognize.

One can, and should, distinguish between two different *sorts* of intentional object for our affective attitudes (whether propositional or otherwise). On the one hand there are the appraisals that provide the inputs to one's valuational systems, most of which are located in subcortical regions of the brain, including the amygdala and basal ganglia. These input-representations are processed for relevance to one's standing goals and values, and a subset of them will issue in an affective response, not only setting in train a variety of physiological (e.g. increased heartbeat) and behavioral (e.g. smiling) changes, but also issuing in positive or negative valence. These appraisals provide what one might think of as the *causal* object, or the *input-side intentional content*, of an affective state.[9]

[9] Note that our basic values aren't stored categorically as structured representational states (as beliefs and memories are). Rather, they are *dispositions* of our evaluative-appraisal mechanisms to respond to a certain class of inputs with a particular sort of affective response.

The valence that results from these input-side appraisals will be directed at, or become associated with, the current object of attention, giving rise to a second set of representational contents. Since these are the ones that guide subsequent planning and intentionally-controlled behavior, they constitute what can be thought of as the *motivational* object, or the *output-side intentional content*, of an affective state. Normally the representations that cause an affective response will at least overlap with the contents one focuses on as a result. (This is because one of the normal effects of affect is to attract attention to its causes; Barrett & Bar, 2009.) But this is not always the case. And there will often be a significant mismatch, with some component of one's affective response deriving from properties other than those one ends up focused on. As a result, what seems desirable or repulsive may be something other than one actually values. That is, the properties that produce the affective response in interaction with standing features of one's value systems may differ from those that one is now motivated to pursue or reject.

The finding appealed to here is one that has emerged out of numerous studies in cognitive science over recent decades (Schwarz & Clore, 1983; Forgas, 1995; Higgins, 1997; Li et al., 2007). Indeed, affective-priming effects are pervasive. Affective states that are caused by one source (a sunny day, finding a nickel in a phone booth, a disgusting environment) can result in affect of the same valence directed towards something else (the overall course of one's life, giving money to charity, the moral wrongness of someone's actions). For example, thirsty people who subliminally view images of happy faces thereafter drink more of a novel beverage, and will pay significantly more for it, than those who subliminally view images of angry faces (Winkielman et al., 2005). The positive valence caused by a happy face primes one to find the drink itself more attractive. Similarly, people who are asked to evaluate the moral wrongness of actions described in short vignettes will rate those actions as morally worse when they complete the task in a disgusting environment, or if they have recently watched a disgusting video (Schnall et al., 2008). Background feelings of disgust prime one to find the described actions to be worse. Moreover, the mere fact that an image is harder to process (resulting in what psychologists call "disfluent processing") produces negative valence, and makes the thing or person depicted in the image seem less attractive (Halberstadt et al., 2013; Halberstadt & Winkielman, 2014). One seemingly confuses the unpleasantness of disfluent perception with the badness of the entity being perceived.[10]

Consider, now, how hedonism appears when seen through the lens of the distinction between input-content and output-content. On the input side, of course, a hedonist will have to allow that one's evaluative mechanisms appraise the objects and outcomes in question in light of both learned and evolutionarily-adaptive values. Appraising the properties of the bear in the bushes, for example, leads to fear (and negative valence) because those properties are a potential threat to one's health and life. Likewise, appraising an action as likely to lead to fame and prestige issues in positive valence because the high

[10] Valence-priming is by no means completely transferrable, however. For different emotions can cause one to construe situations differently. Thus anger causes one to feel in control, and hence reduces perceptions of risk, whereas fear has the opposite effect. So although both are negatively valenced, they can have opposite effects on an ancillary task (Lerner et al., 2015). For example, people might be asked either to write about something that once made them really angry (hence inducing anger) or about something that made them afraid. Shortly thereafter, they may be asked to rate the riskiness of some venture or outcome. Those primed with anger will give lower estimates of risk, whereas those primed with fear will give heightened estimates.

regard of one's community has always been adaptive among group-living humans. But if the valence properties in question are merely qualitative feelings, then what one values on the output side, and chooses in the light of, is just that one should have pleasant experiences and avoid unpleasant ones.

Put differently: on the input side there must generally be world-directed evaluations of actual or potential states of affairs. It is the sight of the chocolate cake that causes positive valence at the thought of eating it, and it is the death of the loved one that causes grief. And even where affect is produced by internal events such as thoughts, memories, or imagination, it is the externally-focused *contents* of those mental events that are received as input by one's evaluative systems and processed to result in affect. It is *what* one thinks about or *what* one imagines that causes an affective response. But on the output side (for a hedonic theorist) the only values are self-directed hedonic ones. There is therefore a stark mismatch between what *one's evaluative systems* care about and respond to and what *the person* cares about and responds to.

This sort of systematic mismatch is quite puzzling, especially from an evolutionary perspective. For evolution couldn't care less about how one feels. What matters is surviving and having descendants, not feeling good. It would therefore be surprising that people should have evolved a decision-making architecture that pays attention only to the prospect of good and bad experiences, requiring a distinct set of beliefs about the objects and events that are likely to issue in such experiences. One might have expected, in contrast, that evolution would have utilized the same representational resources on the input side (when creating affect) as are thereafter employed in one's decision-making about what worldly options and actions to pursue. Since it is the world that is evaluated on the input side, one might have expected that it would be the world that would acquire value as a result.[11]

One can imagine a hedonic theorist responding that perhaps the decision-making architecture *had* to be structured hedonically because mere information about the relevant features of the environment wouldn't be intrinsically motivating. But this reply fails in the absence of further special pleading. For something would have had to wire up hedonic feelings to bias decision-making directly. And it is hard to see a reason why it would have been any more difficult to wire up nonconceptual representations of value to bias decision-making in the same way. So the hedonic account of valence remains evolutionarily puzzling.

The representational account of valence faces no such difficulty, of course. On the input side the bear is appraised as dangerous, and as a result one sees the bear's presence as bad. Likewise, when prestige is appraised as valuable, one comes to see the state of being famous as (nonconceptually) good. One's input-side and output-side values are thus of the same world-directed kind, even if there is often only a partial overlap between

[11] In contrast, there is nothing especially puzzling about the mismatches that result from affective priming. For these would seem to be by-products of the summative function of valence. It is an important part of the design of the system that one's affective mechanisms should be capable of automatically integrating valence from many different sources even if (on occasion) those sources prove to be unrelated to one another. In addition, some have argued that the build-up of valence that results in good or bad moods has the function of speeding up evaluative learning. A good mood will boost the positive valence deriving from good things higher than would otherwise be expected, thus strengthening the evaluative-learning signal and allowing one to adjust to changes in the distribution of goods faster than would be possible through evaluative conditioning alone (Eldar et al., 2016).

the properties that enter into the evaluation and those one comes to see as good as a result. (Note that this sort of lack of overlap is one that a hedonic theorist must acknowledge as well.) It is thus the same kinds of adaptively-relevant properties and events that provide the input-content for a valenced response that also figure on the output-side in the agent's practical reasoning and decision-making.

It is worth noting that, when combined with the ubiquity of valence-priming, the representational account of valence may also prove capable of explaining why hedonism should be so perennially attractive. This is because whenever one experiences a world-directed affective state one *can* transform it into a *self*-directed one just by redirecting one's attention to oneself. Although it is, in the first instance, the presence of the bear that seems bad, if one shifts attention to one's experiences while afraid then those, too (or rather instead), will be seen as bad. Similarly, when in a state of grief, it is the loss of the loved one that is seen as bad in the first instance. But if one shifts attention to oneself, then one's experiences (including of one's bodily state of lassitude) will also come to be experienced as bad. When one attends to oneself, it is the experience of grieving that is represented as bad, rather than the loss of the loved one. In every case, then, someone tempted by or already convinced of the truth of hedonism can find seemingly-confirming evidence just by directing attention toward the self.

Indeed, it may be that some people really *are* hedonists (at least some of the time), because their practical reasoning has a habitual self-focus. Certainly there are measurable differences between people in the extent to which they pay attention to their feelings when responding to stimuli (Gasper & Clore, 2000). It may be that some people, when envisaging a possibility and deciding whether to pursue it, automatically switch attention to their own experiences while envisaging that possibility. As a result, the valence produced by the possibility in question will prime similar valence directed at the self. But it seems unlikely that this sort of habitual self-focus should be the general case. Since the decision-making process almost always begins by considering and appraising worldly possibilities of one sort or another, and since one of the roles of valence is to attract attention to its causes, it seems much more likely that attention will retain its worldly focus for most people, and that it should thus be the worldly action or outcome that is represented as good or bad.

4. Valence and Consciousness

It is obvious that valence must (sometimes) be *access*-conscious—that is to say, available to be reported, to create memories, and to inform decision-making. For the latter, in particular, constitutes its common-currency role. Moreover, it seems, on the face of it, that valence is also *phenomenally* conscious. As we noted earlier, there is something it is *like* to be in agonizing pain, and what it is like seems not to reduce to the phenomenally-conscious status of the sensory component of pain. For the phenomenal quality of one's pain experience is surely quite different if one has undergone anterior cingulotomy, and one's pain sensations are no longer hurtful. However, much else will also be different in this case. If one is no longer bothered by one's pains then one won't tense one's muscles and struggle for relief, which will lead to changes in one's phenomenally-conscious experience. How can we tell, therefore, whether valence makes a *constitutive* rather than a *causal* contribution to phenomenal experience? How can we tell whether valence is phenomenally conscious in its own right?

The most fruitful way of tackling this question, in my view, is to assume that phenomenal consciousness is whatever gives rise to the so-called "hard problem" (Chalmers, 1996), or at least the appearance thereof (Carruthers & Veillet, 2011).[12] That is, phenomenally conscious states are ones for which zombie thought-experiments and other related arguments can sensibly be offered, as well as being states that seem especially ineffable. Arguably this restricts phenomenal consciousness to states with nonconceptual content; and all states with such content that are access-conscious (that is, available to their subjects to report, remember, and enter into planning and decision making) will qualify as phenomenally conscious. For example, one can imagine a being that is physically and behaviorally just like oneself, as well as sharing one's internal functional organization, but for whom there is nothing it is *like* to smell Colombian coffee; and similarly one would be hard pressed to say anything remotely informative about what that smell is like. But one cannot likewise imagine that such being would be incapable of *thinking* (conceptually, propositionally) about Colombian coffee.

If nonconceptual access-conscious content is what constitutes (or at least correlates with) phenomenally conscious experience, as I suggest, then valence will make a constitutive contribution to the phenomenal properties of experience. In support of this claim, note that one would be hard pressed to tell someone who had never felt pain (who has congenital pain asymbolia, for example) what the *hurtfulness* quality of pain is like.[13] Moreover, it seems that one can imagine a being who is like oneself in all respects (including that he groans and cries out when in pain, and wants his pains to cease), but for whom there is *nothing* it is like to be hurt by pain. Furthermore, one can run the equivalent of Jackson's color-deprived-Mary thought experiment (Jackson, 1982, 1986). Imagine someone with congenital pain asymbolia (who feels pain sensations but has never been bothered by them) who becomes a famous psychologist and neuroscientist. She learns everything there is to know about pain (its causes, functions, and physical realization). Surely, one might think, she would nevertheless learn something *new* if she were to be cured of her asymbolia and became capable of being hurt by her pains for the first time. (Compare how color-deprived Mary is supposed to learn something new about color vision when leaving her black-and-white room for the first time.) She might exclaim, "So *this* is what the hurtfulness of pain is like!"

Each of the two views of valence under consideration in this paper can agree that valence makes a constitutive contribution to the phenomenal properties of experience (at least when combined with other assumptions). Put differently, both imply that valence is itself phenomenally conscious (or at any rate that it *can* be; I will return to this point shortly). This is obviously true of the hedonic account. For on this view valence just *is* a qualitative, intrinsic, non-relational property of experience in virtue of the presence of which an experience is felt to be good or bad, welcome or unwelcome. And on this view it is obvious why there is a difference in someone's phenomenal experience of chronic

---

[12] I should stress that while I endorse using the "hard-problem" thought-experiments as a criterion of phenomenal consciousness, this is not to say that those thought-experiments demonstrate that consciousness cannot be reductively explained. Nor do I think that they support the existence of irreducible felt qualities, or *qualia*. That is another topic entirely.

[13] Strictly speaking one needs to imagine a more extreme case of someone who is incapable of any form of negative valence. For if negative valence is a single thing, the same across all forms of negative affective state, as I have suggested, then one *can* tell people with congenital pain asymbolia what the hurtfulness of pain is like. One can tell them that it is like the hurtfulness of grief.

pain before and after undergoing anterior cingulotomy. For following the operation there will be a phenomenal property of the person's experience that is altogether absent, namely the qualitative character that constitutes negative valence.

The representational account of valence entails a similar conclusion, however, at least when combined with some kind of representational theory of consciousness (Tye, 1995). If phenomenal consciousness reduces to access-conscious nonconceptual representation, as Tye and many others believe, then it seems obvious that valence (when access-conscious) will be phenomenally conscious. For valence is a nonconceptual (albeit amodal and nonsensory) representation of value, on this account. Moreover, it will be obvious why valence should seem ineffable, too. For saying that one's pain is *bad* (thereby deploying the concept BAD) will seem wholly inadequate to capture the fine-grained badness that figures in one's experience.

While both accounts of valence can justify the claim that valence can be phenomenally conscious, the hedonic account (and only the hedonic account) seems to entail that valence is *always* phenomenally conscious. At least, this will be true unless one can somehow make sense of the idea of unconscious qualitative character. While some philosophers have speculated that there might be qualitative, phenomenally-conscious, experiences that nevertheless fail to be access-conscious (Block, 1995, 2002), most find this idea hard to accept.[14] It will be equally hard to accept that there can be unconscious forms of valence, then, on a hedonic account of the latter.

In contrast, there need be nothing intrinsically conscious about nonconceptual representations of value. This provides some further reason to accept the representational account, since cognitive scientists now routinely assume that valence can be involved unconsciously at many different levels of processing in the brain, helping to determine the relevance of stimuli for further processing, for example, and helping to select the items that become targets of top–down attention (Barrett & Bar, 2009; Lebrecht et al., 2012; Pessoa, 2013). Recall that we are assuming that valence forms a natural-psychological kind. Then the fact that the relevant science treats valence as a property that can be unconscious as well as conscious provides good reason to prefer an account that coheres with this construal of its nature.

Note, now, that if a hedonic account of valence is correct then theories of phenomenal consciousness that attempt to reduce the latter to representational content of one sort or another will be in trouble. For the intrinsic qualities in question represent nothing beyond themselves. Those of us attracted by representational theories of phenomenal consciousness are thus provided with yet further reason for rejecting a hedonic account of valence. However, some have sought to reverse the direction of this argument, appealing to the felt properties of affective states to mount a direct challenge to representational theories of consciousness, thereby providing an indirect argument *in support of* a hedonic account (Aydede & Fulkerson, 2014; Kind, 2014). Our final task is to evaluate the strength of this argument.

What should a reductive representationalist about consciousness say about the hurtful character of pain? Perhaps one can become convinced that the sensory aspect of pain

[14] Strictly, while Block (1995) says that phenomenal consciousness can be inaccessible, Block (2002) weakens the claim. Phenomenally conscious experiences are said to be access*ible* but often unaccessed. Note that what is at stake when cognitive scientists talk of unconscious valence, however, are representations that are wholly inaccessible to their subjects. So even Block would deny that such states are phenomenally conscious.

represents tissue damage, as Tye (2006) argues. But what natural property is represented by pain's negative valence? One might think it unlikely that there is any natural property of the world that corresponds to negative valence, in part because valence is a *valuational* property (Aydede, 2006). Cutter & Tye (2011) respond to this challenge while assuming a tracking account of representational content in general by arguing that the property tracked by the hurtfulness of pain is that of *harmfulness*. They claim that something is bad for one (as pains generally are) if and only if it is apt to harm one. And harmfulness is surely a natural, biologically explicable, property of events. In which case the *badness* component of pain can be reduced to representations of a naturally-occurring property, after all.

There are a number of things wrong with this reply to Aydede's challenge, however. One is that negative valence is intrinsically motivating, whereas representing something as harmful is not. The account is thus too narrowly cognitive in character, and loses sight of the affective nature of pain.[15] Another problem is that the account fails to generalize to other forms of negatively valenced state. If negative valence is a single thing, as cognitive science seems to suggest, manifested in affective states of many different kinds, then an account of valence should be equivalently general. But it seems implausible that the negative valence present in a sad mood should represent harm. For moods can occur for no reason, without there being any identifiable source of harm, no matter how analogically harm is construed.

If valence is best understood as a nonconceptual representation of goodness or badness, as I have suggested, then we have an account of the appropriate generality. But what, then, are the prospects for a reductive account of the content of these representations? Much may depend on the reductive framework adopted for understanding representational content in general. If the account is an externalist / informational one, purporting to reduce intentional content to the natural properties *tracked* by our representations, or the properties those representations carry *information* about (Fodor, 1990; Dretske, 1995; Tye, 1995; Cutter & Tye, 2011), then one might think that the prospects are not good. For it seems unlikely that there is any natural property that is tracked as *good* or *bad* across the full range of affective states. So while the representational account of valence is consistent with representational theories of consciousness broadly construed, it nevertheless constitutes a challenge to externalist, informational, forms of reductive representationalism about consciousness. This is because nonconceptual representations of goodness and badness presumably do not carry information about, nor causally co-vary with, nor have the proper function of indicating, the presence of goodness and badness in the natural world. Rather, they depend on the partly-innate / partly-learned structure of human and animal reward and valuational systems.

It might be possible to rescue a tracking account of the natural properties represented by valence with a heavy dose of teleology. One might say, for example, that negative valence represents the property of being *maladaptive* while positive valence represents the property of being *adaptive*, provided that one's valuational mechanisms are performing their proper functions. This would enable us to explain what is represented by positive affect at the thought of being famous, given that prestige is generally adaptive. But at the same time we could explain fear of open spaces (which is presumably *not*

[15] In response, however, it might be postulated that there is a distinctive kind of nonconceptual representation of harmfulness that *is* intrinsically motivating.

adaptive) as resulting from some sort of malfunction in one's valuational systems. The account would probably also need to be relativized to the environment of evolutionary adaptation to explain the content represented by positive valence at the taste of sweet, fatty, and salty foods. For strong positive evaluations of these properties were adaptive for our hunter–gatherer forebears, but are no longer so for us.

I will leave it to others to evaluate the prospects for some such externalist reductive account. For this is not my preferred approach to naturalizing intentional content. And even if no such account is successful, this need not mean that one has to retreat toward qualia-realism and property-dualism; nor should one embrace the hedonic account of valence as a result. For reductive forms of functional and/or inferential-role semantics may have no difficulty in accounting for the content of such value representations. Moreover, there also remains the possibility of *non*-reductive naturalism about content, where content is treated as a conceptual primitive (just as it is in much of cognitive science), and where the reality of intentional properties is licensed by the central role they occupy —and are likely to continue to occupy—in successful cognitive-scientific theories (Botterill & Carruthers, 1999; Burge, 2010).

If valence is representational, however, then it must be possible to characterize its correctness-conditions. The hedonic account need assume no such burden, of course. Hedonic feelings may get caused in unusual ways, and may issue in maladaptive actions of one sort or another, but in themselves they are neither correct nor incorrect. For they represent nothing beyond themselves. But if valence represents value, then it must be capable of representing value both correctly and incorrectly. I cannot hope to address this issue adequately here. But one minimalist possibility would be that valence represents value correctly just in case it reflects the underlying dispositional properties of one's affect-generating evaluative mechanisms. At a first pass, one might say that valence of degree *v* directed at an object *o* correctly represents the value of *o* just in case nothing other than *o* contributed to *v*. This would entail that the valence produced in any of the valence-priming experiments discussed earlier *incorrectly* represents the subjective value of the object or event attended to. For in these cases part of the valence is caused, not by the latter, but rather by the disgusting environment, or the sad story one has just written about, or whatever. Whether such an account can be made to work must be left for another time. Here I claim only that it isn't obvious that the correctness-condition burden cannot be adequately discharged.

5. Conclusion

My main goal in this paper has been to put valence, and the question of its nature, firmly on the philosophical map. I have suggested that valence should be given a unitary account across all types of affective state, and that valence may underlie all intentional action. I have argued (albeit non-demonstrably) that the best such account is that valence is an intrinsically motivating nonconceptual representation of goodness or badness, contrasting this with the competing hedonic account along a number of different dimensions. In summary:

(1) The representational account fits the world-directed phenomenology of most affective states better than does the hedonic one.

(2) The hedonic account suggests that only creatures capable of representing their own experiences can engage in valence-based decision-making, whereas the representational account does not.

(3) The hedonic account has difficulty explaining how people can rationally sacrifice future hedonic benefits for something they value. The representational account, in contrast, has no difficulty with such cases. Moreover, it can perhaps explain away the appeal of motivational hedonism, given the pervasive power of valence-priming.

(4) The hedonic account is evolutionarily puzzling, given a systematic mismatch between the worldly inputs to affective processing and the hedonic output.

(5) While both accounts can agree that valence is phenomenally conscious, the hedonic account seems to entail (incorrectly) that valence is *always* phenomenally conscious.

(6) While some theories of representation in general, or of what is represented by the hurtfulness of pain in particular, may have difficulty accounting for the phenomenal content of negative valence, there is no reason to think that representational accounts of valence, as such, are in trouble. Moreover, such theories cohere better with representational theories of consciousness.

Whatever one might think of the details of my arguments, however, my larger conclusion is that philosophers should pay much more attention to the nature of valence and the empirical literature surrounding it. For arguably valence is not only a unitary natural-psychological kind, but one that provides the foundations for all human decision making.[16]

References

Aydede, M. (2006). The main difficulty with pain: Commentary on Tye. In M. Aydede (ed.), *Pain: New Essays on its Nature and the Methodology of its Study*, Cambridge MA: MIT Press.

—— (2014). How to unify theories of sensory pleasure: An adverbialist proposal. *Review of Philosophy and Psychology*, 5, 119–133.

Aydede, M. & Fulkerson, M. (2014). Affect: Representationalists' headache. *Philosophical Studies*, 170, 175–198.

Barrett, L. & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B*, 364, 1325–34.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, 86, 201–21.

Bartra, O., McGuire, J., & Kable, J. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, 76, 412–427.

Bechara, A., Damasio, H., & Damasio, A. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10, 295–307.

Berridge, K. & Kringelbach, M. (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology*, 199, 457–480.

Bird, C. & Emery, N. (2009a). Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proceedings of the National Academy of Sciences*, 106, 10370–10375.

[16] I am grateful to Luca Barlassina, Dan Moller, and an anonymous referee for comments on earlier drafts of this paper.

—— (2009b). Rooks use stones to raise the water level to reach a floating worm. *Current Biology*, 19, 1410–1414.
Block, N. (1995). A confusion about the function of consciousness. *Behavioral and Brain Sciences*, 18, 227–247.
—— (2002). The harder problem of consciousness. *The Journal of Philosophy*, 99, 1–35.
Botterill, G. & Carruthers, P. (1999). *The Philosophy of Psychology*. Cambridge: Cambridge University Press.
Bratman, M. (1987). *Intentions, Plans, and Practical Reason*. Cambridge MA: Harvard University Press.
—— (1999). *Faces of Intention*. Cambridge: Cambridge University Press.
Buckner, R. (2010). The role of the hippocampus in prediction and imagination. *Annual Review of Psychology*, 61, 27–48.
Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press.
Carruthers, P. & Veillet, B. (2011). The case against cognitive phenomenology. In T. Bayne & M. Montague (eds.), *Cognitive Phenomenology*, NY: Oxford University Press.
Chalmers, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press.
Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204.
Corbetta, M., Patel, G., & Shulman, G. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, 58, 306–324.
Cutter, B. & Tye, M. (2011). Tracking representationalism and the painfulness of pain. *Philosophical Issues*, 21, 90–109.
Damasio, A. (1994). *Descartes' Error*. London: Papermac.
Dretske, F. (1995). *Naturalizing the Mind*. Cambridge MA: MIT Press.
Durso, G., Luttrell, A., & Way, B. (2015). Over-the-counter relief from pains and pleasures alike: Acetaminophen blunts evaluation sensitivity to both negative and positive stimuli. *Psychological Science*, 26, 750–758.
Eldar, E., Rutledge, R., Dolan, R., & Niv, Y. (2016). Mood as representation of momentum. *Trends in Cognitive Sciences*, 20, 15–24.
Ellingsen, D-M., Wessberg, J., Eikemo, M., Liljencrantz, J., Endestad, T., Olausson, H., & Leknes, S. (2013). Placebo improves pleasure and pain through opposite modulation of sensory processing. *Proceedings of the National Academy of Sciences*, 110, 17993–17998.
Ellsworth, P. & Scherer, K. (2003). Appraisal processes in emotion. In R. Davidson, K. Scherer, & H. Goldsmith (eds.), *Handbook of Affective Sciences*. New York: Oxford University Press.
FitzGerald, T., Seymour, B., & Dolan, R. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. *Journal of Neuroscience*, 29, 8388–8395.
Fodor, J. (1990). *A Theory of Content and Other Essays*. Cambridge MA: MIT Press.
Forgas, J. (1995). Mood and judgment. *Psychological Bulletin*, 117, 39–66.
Gasper, K. & Clore, G. (2000). Do you have to pay attention to your feelings to be influenced by them? *Personality and Social Psychology Bulletin*, 26, 698–711.
Gilbert, D. & Wilson, T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science*, 14, 131–134.
—— (2007). Prospection: Experiencing the future. *Science*, 317, 1351–1354.

Griffiths, P. (1997). *What Emotions Really Are*. Chicago: University of Chicago Press.
Gross, J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26, 1–26.
Halberstadt, J. & Winkielman, P. (2014). Easy on the eyes, or hard to categorize: Classification difficulty decreases the appeal of facial blends. *Journal of Experimental Social Psychology*, 50, 175–183.
Halberstadt, J., Pecher, D., Zeelenberg, R., Wai, L., & Winkielman, P. (2013). Two faces of attractiveness: Making beauty in averageness appear and reverse. *Psychological Science*, 24, 2343–2346.
Hanus, D., Mendes, N., Tennie, C., & Call, J. (2011). Comparing the performances of apes (Gorilla gorilla, pan troglodytes, Pongo pymaeus) and human children (Homo sapiens) in the floating peanut task. *PLoS One*, 6, e19555.
Harmon-Jones, E. & Mills, J. (1999). *Cognitive Dissonance*. Ann Arbor: Braun-Brumfield.
Izard, V., Sann, C., Spelke, E., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, 106, 10382–10385.
Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32, 127–136.
—— (1986). What Mary didn't know. *Journal of Philosophy*, 83, 291–295.
Johansson, P., Hall, L., Tärning, B., Sikström, S., & Chater, N. (2014). Choice blindness and preference change: You will like this paper better if you (believe you) chose to read it! *Journal of Behavioral Decision Making*, 27, 281–289.
Jordan, K., MacLean, E., & Brannon, E. (2008). Monkeys match and tally quantities across senses. *Cognition*, 108, 617–625.
Kind, A. (2014). The case against representationalism about moods. In U. Krigel (ed.), *Current Controversies in Philosophy of Mind*, London: Routledge.
Lambie, J. & Marcel, A. (2002). Consciousness and the varieties of emotion experience. *Psychological Review*, 109, 219–259.
Lebrecht, S., Bar, M., Barrett, L., & Tarr, M. (2012). Micro-valences: perceiving affective valence in everyday objects. *Frontiers in Psychology*, 3: #107.
LeDoux, J. (2012). Rethinking the emotional brain. *Neuron*, 73, 653–676.
Leknes, S. & Tracey, I. (2008). A common neurobiology for pain and pleasure. *Nature Reviews Neuroscience*, 9, 314–320.
Leknes, S., Berna, C., Lee, M., Snyder, G., Biele, G., & Tracey, I. (2013). The importance of context: When relative relief renders pain pleasant. *Pain*, 154, 402–410.
Lerner, J., Li, Y., Valdesolo, P., & Kassam, K. (2015). Emotion and decision making. *Annual Review of Psychology*, 66: #33.
Levy, D. & Glimcher, P. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, 22, 1027–1038.
Li, W., Moallem, I., Paller, K., & Gottfried, J. (2007). Subliminal smells can guide social preferences. *Psychological Science*, 18, 1044–1049.
Lieberman, M. (2013). *Social: Why our brains are wired to connect*. New York: Crown Publishers.
Lieberman, M. & Eisenberger, N. (2009). Pains and pleasures of social life. *Science*, 323, 890–891.
Lieberman, M., Ochsner, K., Gilbert, D., & Schacter, D. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*, 12, 135–140.

Lisman, J. & Sternberg, E. (2013). Habit and nonhabit systems for unconscious and conscious behavior: Implications for multitasking. *Journal of Cognitive Neuroscience*, 25, 273–283.

McDowell, J. (1994). *Mind and World*. Cambridge, MA: Harvard University Press.

Moller, D. (2011). Anticipated emotions and emotional valence. *Philosopher's Imprint*, 11, 1–16.

Panichello, M., Cheung, O., & Bar, M. (2013). Predictive feedback and conscious visual experience. *Frontiers in Psychology*, 3: #620.

Pessoa, L. (2013). *The Cognitive-Emotional Brain*. Cambridge MA: MIT Press.

Plassmann, H., O'Doherty, J., & Rangel, A. (2010). Appetitive and aversive goal values are encoded in medial orbitofrontal cortex at the time of decision making. *Journal of Neuroscience*, 30, 10799–10808.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, 105, 1050–1054.

Prinz, J. (2004). *Gut Reactions*. Oxford: Oxford University Press.

Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology*, 67, 525–539.

Rolls, E. (1999). *The Brain and Emotion*. Oxford: Oxford University Press.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.

—— (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145–172.

Schnall, S., Haidt, J., Clore, G., & Jordon, A. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34, 1096–1109.

Schroeder, T. (2004). *Three Faces of Desire*. Oxford: Oxford University Press.

Schwarz, N. & Clore, G. (1983). Mood, misattribution, and judgments of well-being: Informative affective states. *Journal of Personality and Social Psychology*, 45, 513–523.

Seeley, W., Menon, V., Schatzberg, A., Keller, J., Glover, G., Kenna, H., Reiss, A., & Greicius, M. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27, 2349–2356.

Seligman, M., Railton, P., Baumeister, R., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science*, 8, 119–141.

Sharot, T., De Martino, B., & Dolan, R. (2009). How choice reveals and shapes expected hedonic outcome. *Journal of Neuroscience*, 29, 3760–3765.

Sharot, T., Fleming, S., Yu, X., Koster, R., & Dolan, R. (2012). Is choice-induced preference change long lasting? *Psychological Science*, 23, 1123–1129.

Sober, E. & Wilson, D. (1999). *Unto Others: The evolution and psychology of unselfish behavior*. Cambridge MA: Harvard University Press.

Solomon, R. (1984). *The Passions*. New York: Doubleday.

Sripada, C. (2007). Adaptationism, culture and the malleability of human nature. In P. Carruthers, S. Laurence, & S. Stich (eds.), *Innateness and the Structure of the Mind*, Vol. III. Oxford: Oxford University Press.

Stich, S., Doris, J., & Roedder, E. (2010). Altruism. In J. Doris & The Moral Psychology Research Group (eds.), *The Moral Psychology Handbook*, Oxford: Oxford University Press.

Taylor, A., Elliffe, D., Hunt, G., & Gray, R. (2010). Complex cognition and behavioral innovation in New Caledonian crows. *Proceedings of the Royal Society B: Biological Sciences*, 277, 2637–43.

Tye, M. (1995). *Ten Problems of Consciousness*. Cambridge MA: MIT Press.

—— (2006). Another look at representationalism about pain. In M. Aydede (ed.), *Pain: New Essays on its Nature and the Methodology of its Study*, Cambridge MA: MIT Press.

Völter, C. & Call, J. (2014). Younger apes and human children plan their moves in a maze task. *Cognition*, 130, 186–203.

Wager, T. (2005). The neural bases of placebo effects in pain. *Current Directions in Psychological Science*, 14, 175–179.

Wilkinson, H., Davidson, K., & Davidson, R. (1999). Bilateral anterior cingulotomy for chronic noncancer pain. *Neurosurgery*, 45, 1129–1136.

Wilson, T. (2010). *Redirect: The surprising science of psychological change*. London: Penguin.

Winkielman, P., Berridge, K., & Wilbarger, J. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, 31, 121–135.