

Reasons and Theories of Sensory Affect

Murat Aydede (UBC) & Matthew Fulkerson (UCSD)

1 Introduction

Some sensory experiences are pleasant, some unpleasant. This is a truism. But understanding what makes these experiences pleasant and unpleasant is not an easy job. Various difficulties and puzzles arise as soon as we start theorizing. There are many philosophical theories on offer that give different accounts of the positive or negative affective valences of sensory experiences. In this paper, we will look at the current state of art in the philosophy of mind, present the main contenders, and then critically compare and contrast them. In particular, we want to examine how they handle the *reason-giving power* of affective states. We will look into two representationalist proposals and a functionalist proposal, and argue that, contrary to their own advertisements, the representationalist proposals don't adequately account for why and how sensory affect can motivate, rationalize, and justify subsequent behavior and intentional mental activity. We will show that our own functionalist proposal does a much better job in this regard, and that when the representationalist proposals are modified to do a better job, they may fare better not because of their representationalist credentials but due to their functionalist ones.

2 Preliminary Taxonomy

We will concentrate on those sensory episodes that are unpleasant (including the painfulness of sensory pains) — although most of what we say, when modified appropriately, will be applicable to positive, pleasant sensory episodes. Arguably, some pains — those of pain asymbolics, for instance — are not painful or unpleasant. While we do not discuss such states here, our unifying framework offers a robust means of explaining away such unusual cases.

Our target then are sensory pains understood as sensory experiences that are unpleasant. What is unpleasant about them seems initially to be the experiences themselves. Keeping this in mind we have the following truism:

(T) Sensory pains are unpleasant experiences.

The job, then, is to specify the following schema as best as one can:

(SCHEMA) Sensory experiences (e.g., pains) are unpleasant if, and only if, ...

Depending on how the ellipsis is filled, theories can be reductive or non-reductive. So, for instance, consider the class of popular *felt-quality* views of sensory pain:

Felt-quality: Sensory (pain) experiences are unpleasant (painful) if, and only if, they have a common phenomenological aspect distinctive of unpleasant (pain) experiences.

As stated, this is a non-reductive view that can be generalized to cover all affective valence — positive or negative. There are different versions of it in the literature, including *hedonic-tone* views and *distinctive-feeling* views.¹ Typically, a defender of these views resists reductive proposals by claiming that the phenomenological aspect in question is a primitive occurrence — either a distinctive phenomenal feature common to all pleasant or unpleasant sensations, or a somewhat varying phenomenal occurrence in a hedonic dimension that cannot be further analyzed. We will not consider such non-reductive views in what follows (but see Aydede 2014).

In principle, a felt-quality theorist could, of course, go on to make a reductive claim by identifying the phenomenological aspect peculiar to sensory affect with a certain sort of neurophysiological event, or with a certain sort of causal role that the sensory experience plays in the larger mental economy of the subject, or with a certain kind of representation or representational content, etc. It may be that these various reductive options remove the appeal of the felt-quality views. Arguably, once a reductive claim is made, all or most of the explanatory work may be done by the reductive component of the theory.

There are two broad approaches one may take in offering a reductive account. First, one could take the task of filling in the ellipsis in SCHEMA to be a matter of supplying a *conceptual truth* derived via the conceptual analysis of pleasure and pain. For instance, the influential *attitudinal* (or, *externalist*) theories developed in value theory and moral psychology can be taken this way. According to such theories, a pleasant sensation is a composite state, composed of a target sensation and some mental (usually, conative) attitude directed toward that sensation, such as desiring, wanting, preferring, or liking the sensation in some sort of way.² The motivation for this view is that the very concept of

¹ For distinctive feeling views see (among others) Moore (1903/1993: 12), Brink (1989: 221) and Bramble (2011). For the hedonic tone view, Broad (1930), Duncker (1941), Kagan (1992), Sprigge (2000), Crisp (2006), Smuts (2011), and Labukt (2012), among others.

² Defenders of attitudinal views include Alston (1968), Davis (1981, 1982), Brandt (1979), Feldman (1997, 2004), Heathwood (2007), among others. Parfit (2011) and Brady (2013) develop the attitudinal theory in terms of liking (or disliking) a sensation, where this hedonic attitude is taken to be different from desiring in

sensory pleasure or painfulness seems to require that such experiences be essentially connected to other mental states. For instance, one may argue that the very concept of a painful pain that is not disliked or intrinsically motivating is incoherent. Such a state would, according to these theorists, not count as painful at all. Such views are not confined to those in moral psychology; similar proposals have also been made by philosophers of mind, often in the spirit of analytical functionalism. These sorts of views analyze sensory pleasure or painfulness in terms of how these sensations need to be causally related to personal level mental attitudes.³

The second reductive approach one could take is to complete SCHEMA in such a way that the resulting proposal would be *a posteriori*. Rather than filling in the ellipsis merely on the basis of conceptual analysis, we investigate actual states of pain and pleasure and discover their underlying nature. This investigation can involve some speculative theory construction, but it is intended to be scientifically-informed and sensitive to empirical facts. In the remainder of this paper we will focus on three such proposals. Two of these theories are broadly representationalist: *evaluativism* and *imperativism*. Both attempt to reduce sensory affect to a certain kind of content (as the names imply, to evaluative and imperative contents, respectively). The third alternative is the account we defend: *psychofunctionalism*. Rather than reduce sensory affect to a certain kind of content, our view identifies sensory affect (painfulness, pleasantness) with the causal and computational roles typical of sensory affect.

Before we consider these views in more detail, we would like to mention briefly another way of carving out the classificatory scheme of accounts of sensory affect.

We started with the claim that (T) is a truism. What makes (T) a truism is that it seems to be an obvious *phenomenological* truth. Anyone who has the relevant concepts and is capable of making mundane introspective observations could, it seems, confirm it instantly without any experimental procedure. Here we understand “phenomenology” to be the relatively robust phenomenology of sensory experiences for which the question of what it’s like to have this or that sensory experience has (usually) a fairly straightforward personal-level answer — even though it may be very difficult or even impossible to express it in public languages. Indeed, to some extent, felt-quality views may be taken as expressing the default commonsense view that sensory affect is phenomenologically real and robust.

Many have contested felt-quality views on precisely this point: critics have argued that (T), if it’s a truth, is not a *phenomenological* truth in any robust sense. Indeed, many

that it is supposed to be not responsive to reasons. It is difficult to classify Sidgwick (1907): he offers what appears to be an attitudinal theory but he uses ‘desirable’ rather than ‘desired’ (1907/1981: 127).

³ See, for instance, Armstrong (1962, 1968), Pitcher (1970), Tye (1995), Hall (1989), Kahane (2009). Robinson (2006) gives an account of sensory pleasure that is hard to classify but seems closer to the attitudinal theories as he takes the pleasantness to essentially involve intentional and evaluative directedness toward sensations.

defenders of attitudinal views (especially in value theory) take their view to be in competition with felt-quality views and do so in a way that denies (T) as expressing an introspectively available phenomenological truth. For the most part, they are motivated by what is sometimes known as the *heterogeneity problem* — the problem of finding a single phenomenal quality common to all and only pleasant sensations (or unpleasant sensations as the case may be). Once this problem is made clear, (T) as a phenomenological claim seems in jeopardy, and it seems more plausible that something else — an attitude, say — is the common factor between headaches, sun burns, and cases of nausea.

This suggests another possible taxonomy which contrasts *phenomenal realism* with *phenomenal irrealism* about sensory affect. The debate between these groups is an interesting one, for sure — the phenomenology of sensory affect is poorly understood and is certainly in need of better study, philosophical or otherwise. As things stand now, however, it is difficult to settle the debate stated in these terms as the arguments on both sides tend to rely more or less exclusively on intuition and introspective evidence — the very elements on which the sides disagree. This much is clear: *if* sensory affect has a phenomenological reality, it is not as robust as the *sensory* phenomenology of standard cases of perceptual contact with the extra-mental world.⁴ At any rate, following commonsense and the majority of thinkers in this area, we will assume (at least, a weak version of) phenomenological realism about sensory affect for the rest of this paper and read (T) as stating a phenomenological truth. However, nothing much depends on this for what follows — the discussion could be conducted just as well without much modification under the assumption of irrealism.

3 Reductive Theories of Sensory Affect

Let us now consider our three contemporary *a posteriori* reductive theories of sensory affect.

We will start with the two representational accounts. Representationalism (sometimes *intentionalism*) is a popular view in contemporary philosophy of mind identifying the phenomenal character of a conscious experience with its representational content. In ordinary cases of perception the idea is that the phenomenal character of perceptual experience seems exhausted by the features represented in the experience. As Harman says in a classic passage, “When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experience” (Harman 1990: 39).⁵

⁴ Even this much may not be agreeable to everyone. There are theorists who think that experiential affect is the perception of objective value out there. Perhaps they think that its phenomenology is as robust — whatever phenomenological robustness comes to. It’s hard to tell. See Johnston (2001).

⁵ Defenders of representationalism include (among many others) Tye (1995), Dretske (1995), and Byrne & Tye (2006).

Such standard representational accounts of perceptual experience, while popular, have not been without their critics. And, whatever their general problems, it should be clear that, when it comes to accounting for sensory affect, they have a particular difficulty. For, whereas the sensible features represented in a typical visual experience have plausible objective, physical analogues (colors, shapes, locations, distances, textures, etc.), it's not at all clear what sensible features affective qualities might turn out to be or correspond to. The difficulty with standard descriptive proposals — that is, treating pain as precisely analogous with vision and other perceptual experiences — is that when Eloise feels pain it's both very difficult to say what external quality she might be representing and to deny that she is aware of some intrinsic, hurtful quality to her experience.

Rather than pursue this quite general debate here, we are going to look at two particular representationalist accounts that seek to avoid the problems faced by unsophisticated versions of the view.⁶ Both seek to carve out a new area of logical space by denying continuity between typical perceptual experiences and cases of sensory affect. Instead, they argue that affective and perceptual experiences share a broadly representational structure (both can be reductively explained by appeal to content) but they critically differ in the kind of content they carry.

3.1 The Imperative View

The first such view we will consider is imperativism. This is the view that pains and other motivating states specify commands: they involve sensory content that tells us what to do (or what not to do). Rather than descriptive content, the imperativist believes affective states possess *imperative contents* expressed by these commands. The content constitutive of affective experience is a command to the subject to perform some particular action (or in some cases, to refrain from acting). It is important to understand that, in context, the imperative view arises as a way of saving representationalism (indeed, of saving a strong version of the view) by expanding the possible *moods* that representational content can possess, thereby sidestepping the problems typical descriptive views have with affective states.

The commands adverted to by the imperative view are not taken to be a forced movement of our bodies (it's not a reflex action — as its defenders make clear). Nor is the command specified or explained by a separate attitude like desire, emotion, or belief that accompanies the sensory experience. According to the imperativist, the command is understood as the *imperative representational content* of our affective experiences, and this content in imperative mood is the reductive explanatory base for what it's like to have an affective sensory experience of that kind.⁷

⁶ See, however, Tye (2005), Aydede (2005, 2009a).

⁷ Klein (2007) uses 'intentional' rather than 'representational', reserving the latter for only descriptive/indicative content. We will continue to use 'representational' as having (conceptual or non-

Richard Hall (2008) argues that itches (but also hunger, thirst, and pain) are best understood by their possession of imperative contents:

On my view, itches are experiences with a content, as the intentionalists argue. But the intentional content isn't a representation of some bodily state ... 'Representation' suggests descriptive content but itch experiences don't have descriptive content. Itches don't describe some state of the body at the felt location. Rather the intentional content of an itch is an imperative: 'Scratch!' Not in English, of course, but in sensory code. (2008: 526)

On Hall's view, the imperative content is the motivating and affective part of the experience, and can be understood as part of an overall experience that also contains (some) descriptive contents.

Colin Klein (2007) defends a more austere version of the view, this time explicitly directed at pain:

I argue that pains are exhausted by their content, but that this content is imperative rather than representational. Pains thus command rather than describe. Commanding is still a way of having content, however, and so intentionalism is preserved. (2007: 518)

Klein's view explicitly parallels strong versions of the representational view, only swapping out imperative content for descriptive. Instead of positive imperatives like "scratch!", "drink!" and "eat!", Klein argues that pains involve negative imperatives: "Do not <Action>!" For example, a pain in the foot might have the content: "Do not place weight on the foot!"

More recently, Manolo Martínez (2011) has also defended a version of imperativism for pain. According to Martínez, "The content of a headache is one with satisfaction conditions that we could render more or less as: *that this bodily disturbance is no more!*" (Martínez 2011: 78). He goes on to suggest some advantages this particular formulation has over Hall's and Klein's versions, and to suggest some ways that the view can be naturalized (more on this last point shortly).

While these versions differ, sometimes importantly, in their details, for our purposes these differences don't matter. What is crucial for us is that all versions of the view appeal to a special representational content that is understood in terms of commands. Generalizing the view:

conceptual) propositional content in whatever attitude or mood the representations are deployed, and use 'representationalism' to label the view that all experiences are representational.

Imperativism: *S*'s pain is unpleasant if, and only if, *S* receives a command from her relevant sensory system to the effect that *S* engage (or, stop engaging) in a certain sort of behavior directed to the bodily location where the pain is felt.

A few clarifications are in order before we can assess the general view. The imperativist needs to specify who or what issues the command as well as who receives it. Issued commands are supposed to be representational items with satisfaction conditions. These conditions need to be specified at least in general terms. Similarly, commanding requires a representational medium in which commands are issued. Thus, the format of commands needs to be specified too.

We will assume (as Hall makes explicit) that the representational medium is a sensory code in which our sensory experiences are realized. So the representation in question is experiential (perhaps, non-conceptual with presumed phenomenology) and issued in the imperative mood commanding the subject to do whatever it takes to change the disordered bodily condition that is the cause of the pain. We will follow Martínez (2011, 2013) in taking this command to be something like:

Pain Command: See to it that this bodily disturbance is no more!

The command is about a physical disturbance in or on the subject's body to the effect that the subject take steps to get rid of it. Let's also say that the command is issued by the subject's pain module — a specialized sensory system that monitors and reports on potential bodily disturbances or disorders. This command is issued to the subject and is received when the subject experiences the pain as unpleasant or painful. These glosses will suffice for the moment.

The key question concerning imperative contents for pain concerns how such contents gain their imperative mood. As Hall notes in a footnote:

How can a sensory experience have an imperative mood or force? Although possible answers come to mind (perhaps Mentalese, or more particularly sensory-ese, contains explicit indicators of imperativeness, or perhaps the functional role of the message — where it's from, where it goes, how it's operated on — constitutes its imperativeness), this is a topic for another paper, since it is a general problem for any intentional view of experience. How, after all, does a 'descriptive' representation, say from vision, gets its indicative mood and its declarative or assertive force? (2008: 526)

Note that Hall here openly grants the possibility that the functional role of a representation is essential for its content to be imperative.⁸ This becomes clear when he suggests that this is not a worry since the indicative versions of representationalism face the same problem. This is somewhat misleading, however, for the most plausible accounts of how to cash out indicative contents — various tracking theories of representationalism — are simply not available for typical motivating states (see Aydede and Fulkerson 2014). We will argue below that this problem is not so easily dismissed, and that the only plausible way to cash out imperative content is by invoking some version of (psycho)functionalism.

3.2 Evaluativism

Let's now turn to the other form of representationalism, a view that David Bain calls *evaluativism*.⁹

According to the evaluativist, affective sensory states can be reductively explained by their *evaluative content*. Such content is not purely descriptive (as in ordinary perceptual representational views) nor are they commands as on the imperative view. Instead, the view holds that affective experiences involve contents that make some normative or evaluative claim on the world, usually in terms of something being good or bad for the subject. Again, generalizing away from the details of the various extant proposals (and following Bain's canonical formulation of the view), we can state the essential elements of the view:

Evaluativism: “A subject's being in unpleasant pain consists in his (i) undergoing an experience (the pain) that represents a disturbance of a certain sort, and (ii) that same experience additionally representing the disturbance as bad for him in the bodily sense.” (Bain, 2013: S82)

Again, these representations are experiential and (perhaps) non-conceptual with appropriate phenomenology. It is the second clause that makes this form of representationalism evaluativist. The claim here, unlike that made by the imperativist, is that the mood of this representation is, like those for normal perceptual experiences, indicative. Or at least, it is a representation that is truth-apt, with legitimate truth-conditions. What differs from the standard perception case isn't the *mood* of the content, but rather the nature of the properties represented. Whereas standard perceptual experience puts us into contact with neutral sensible qualities, affective experiences put us into contact with value-laden or

⁸ By invoking Millikan's consumer semantics as a natural avenue for naturalizing imperative content, Martínez also seems to implicitly support the reduction to psychofunctional role — see below.

⁹ Versions of this view have been held by Bain (2013); O'Sullivan & Shroer (2012); and Cutter & Tye (2011, 2014), Tye (2005). An early and influential version of this view seems to be Helm's (2002) *Felt Evaluation* view — although Helm seems to develop and defend his view on *a priori* grounds.

normative qualities. These are still thought to be objective features on most versions of the view (although there is interesting logical space here for a non-naturalist evaluative theory, including a so-called *Edenic*-account — see Chalmers 2006), characterized by what is, as a matter of fact, good for the subject (or, on some accounts, the subject's body). These contents are normative in the sense that they specify, for example, what qualities an agent *should* or *should not* seek to avoid, or by stating which qualities are intrinsically *good* and which *bad*. The main challenge for such a view, as one might expect, is to give a plausible account of the nature of these evaluative contents. Various proposals have been made in the literature, and some of them (i.e., Cutter & Tye 2011) are fairly complex. Our view is that none of the proposals so far introduced have given an adequate, naturalistically satisfying account of the nature of evaluative content. (We should be clear, this does not necessarily stem from an aversion to all normative content, only to the use of such content to explain episodes of affective sensory awareness).

3.3 Psychofunctionalism

Finally we turn to a rough sketch of our preferred psychofunctionalism. Unlike the previous accounts, the psychofunctionalist view does not attempt to reduce the phenomenal character of sensory affect to a special kind of representational *content* or to a special *mood* for such content (this is not to say that the view denies pleasant and unpleasant states involve representations — see below). Instead, the psychofunctionalist can trace their motivation to the idea present in analytic functionalism that pains and other affective sensory states seem to be essentially linked to a suite of behavioral and dispositional reactions. Unlike these earlier forms of functionalism, the psychofunctionalist is not interested in mere conceptual analysis or in giving an account in behaviors and dispositions characterized solely at the personal level. Instead, the view works on the assumption that our best account of the causal influences, computations, and dispositions constitutive of the target mental state will be uncovered in (largely) subpersonal terms by work in the relevant empirical sciences (especially neuroscience and cognitive psychology). The reductive base then may not include causal roles immediately available to introspection or to our naïve conception of the ordinary notions involved. The causal roles involved will be, for the most part, a substantive and theoretically robust *discovery*, couched ultimately in the language of the most reliable empirical sciences.

For our purposes, we can, to some extent, stay neutral on the nature of the functional roles to be uncovered. While any specific view will ultimately have to make some choices here, our defense of the view can also remain silent about more general issues concerning whether the functional roles are purely causal, computational, or teleological; or whether the ontology of functions is process-based or substantival (or both). Indeed, we want to leave open the possibility that psychofunctional roles may be constituted by distinct kinds of functions. Generalizing the view:

Psychofunctionalism: A subject's pain experience is unpleasant if, and only if, the sensory information carried by the pain experience plays a certain causal/psychofunctional role (call this role, CR).¹⁰

As we noted, specifying CR isn't exclusively or even mostly a philosopher's job. Certainly not at the most detailed levels. Nevertheless, it is important to fill in the ellipsis by outlining the kind of causal role we have in mind. Work in recent affective neuroscience suggests that this role consists of a complex processing of incoming sensory information that, among other things,¹¹

- sets interruptible motivational parameters (motivational biasing — "more-of-this" or "less-of-this" or "stop-this" incoming stimulus);
- prepares and primes the effector or psychomotor systems of the organism, providing action-preparedness (motor biasing);
- provides appraisals of the incoming sensory information for its significance for the organism and for its potential for enhancing its behavioral repertoire (epistemic biasing);
- influences action preferences on the basis of the sensory stimuli's informational content for present and future behavior through associative or cognitive learning, habituation, incentive sensitization, etc.;
- provides steady, earmarked input to more centralized concept wielding cognitive and decision-theoretic or conative systems.

¹⁰ The psychofunctionalist, unlike the analytical functionalist, need not specify a theoretical identity or a set of necessary and sufficient conditions on the explanandum right at the beginning. Instead, the theorist can remain neutral about the ultimate character of the theory, and simply proceed with functional analysis. The task is one of reductively explaining a complex mental state by breaking it down into its functional parts, and then showing how those various parts work together to produce the higher-level state (see Cummins 1981). We can decide on the nature of the identities and linking propositions after the spade work has been done, as a kind of mopping up operation. Importantly, while there need not be theoretical identities at the lowest levels (for mundane reasons having to do with multiple realizability and functional constancy), at the higher levels of analysis, where the philosopher works, there will be something like a reductive *a posteriori* identity between the target state and a (higher-level) functional role. It is precisely at this level that we plan to couch our theory. We do this by coordinating person-level and empirical work to define a set of high level roles constitutive (or least, characteristic) of sensory affective states — see *FIGURE* below. See also Clark (2005) who is a fellow psychofunctionalist about the painfulness of pains. Tim Schroeder's (2004) work on pleasure is also generally in the same psychofunctionalist ballpark — although he seems to think of himself as a representationalist.

¹¹ See Berridge (2004), Rolls (2005), and Panksepp (1998). We do not take these bullet points to exhaust CR, nor do we take them, individually or even jointly, to be *very* illuminating. The point is to show intuitively the sort of psychofunctional causal role we have in mind. 'CR' is for now a place holder for whatever the ultimate scientific account of this role will turn out to be. In future work, we hope to fill in the system-level empirical details.

One crucial aspect of this proposal is that the distinctive affective processing is inherently motivational in a broad sense. This leads us to call this style of processing *m-processing*. Keeping in mind that while m-processing is a general term for a heterogeneous class of distinct functional processes and realizers,¹² we can still state a non-trivial explanatory reduction in terms of a theoretical identity by holding that sensory affect *just is* the m-processing of incoming sensory signals. We will say that the affective valence of sensory experiences (including the unpleasantness of pains) consists of the sensory information contained in the experience being m-processed, i.e., being subjected to the kind of processing outlined above and whose details are to be filled by empirical sciences. We believe that the metaphysics of sensory affect is a functional affair and that this is in fact supported both by findings in the sciences but also (to some extent) by the phenomenology of such experiences. Let us elaborate on this a little further.

The functional nature of m-processing is not altogether hidden at the personal level — this is what makes attitudinal theories, analytic functionalism, and in general conceptual reductionism, *prima facie* plausible. The pleasantness of the taste of a ripe strawberry consists literally in our being motivationally (motorically, epistemically, etc.) *biased* toward that very taste *and* toward the strawberry in the very experiencing of that taste. This is how that taste *and* the strawberry are being presented to our consciousness. The pleasantness or the unpleasantness, in this sense, is the “felt evaluation” of that taste and the strawberry (Helm 2002). Clearly, some sort of a “desire-like” state is involved somewhere in the phenomenology of sensory affect. When we have a pleasant sensation such as the taste of the ripe sweet strawberry, there is some sort of a tug that we feel toward the taste, we feel some kind of pull that makes us “want to continue having the sensation”.¹³ Not only that, and more importantly, the pull is also toward the strawberry, the object of our taste experience. But the ordinary intuitive notion of desire isn’t up to capturing this. If there is any sense in which there is a kind of desire involved in affective experiences, it is a *phenomenologically salient experiential desire*. Let us call this *phen-desire*. Just as the sensory informational content of a perceptual experience can be encoded non-conceptually in the sensory code (descriptive sensory phenomenology, call it *phen-believing*),¹⁴ the affective content of the experience can be identified with the way this sensory information

¹² There is growing evidence that affective processing of sensory information is done in both parallel and serial fashion in mutually interacting distinct computational modules. For instance, the mechanisms responsible for computing the reward value of a stimulus, for the utilization of this information for learning, and for the execution of behavior subsequent to this seem to be distinct, and there is strong evidence that they can and do come apart. See, for instance, Pecina *et al.* (2006).

¹³ Brandt (1979: 38) gives this as a definition of sensory pleasure.

¹⁴ Compare to Byrne’s *exing that P* (2009) — roughly *experiential/phenomenal believing that P* that feeds into the mechanisms of ordinary believing under certain conditions. See also Aydede (2000) for a similar explanation of m-processing as, what he calls, ‘desiring*’.

is m-processed for motivational/learning related purposes (affective/conative phenomenology).¹⁵ It is a form of *experientially desiring* (phen-desiring) the object of the experience whose sensible properties are also phenomenologically registered (phen-believed) in the sensory content of the very same experience. So we have the *experiential* (non-conceptual) analogs of the ordinary notions of belief and desire (these latter generally conceived as conceptually structured propositional attitudes).

In the case of normal sensory pains, the experience of pain would involve sensorially encoding the physical disorders in the body while presenting them as the objects of our phen-desires (maybe to the effect that they cease or go away). The phenomenology of this phen-desiring would be tantamount to the pain's being painful, unpleasant — its negative affective character. We would like to note here that the affective qualities such as pleasantness and unpleasantness (the affective phenomenology) primarily qualify the affective *experiences* as loci of information processing, not the things or events this information is about. Phen-desiring may be a sensory desire for the extra-mental objects the sensory information is about, but the phenomenology is primarily the phenomenology of the phen-desiring (\approx m-processing), not of its object. What is pleasant or unpleasant are the experiences, not the objects of those experiences. If there is a sense in which the *objects* of our sensory/affective experiences are pleasant or unpleasant, it is only in a derivative sense: they are apt to cause pleasant or unpleasant experiences. The primary locus of affect is the experience itself.¹⁶ It is important to keep this in mind because it will be crucial in understanding how the experiences themselves can be (intrinsically) good or bad (as opposed to their objects being so) and can thus be the objects of our ordinary desires — we will return to this below.

Thus, even though the metaphysics of sensory affect is psychofunctionalist (\approx negative or positive m-processing), our view has a lot of insights to offer at the personal (phenomenological) level by identifying this m-processing with (roughly) a kind of pro or con experiential desire (phen-desire) that can make the experiences themselves be the objects of our ordinary desires. (See *FIGURE*.)

¹⁵ See Aydede (2014, forthcoming) who likes to explain the phenomenology of this affective modification of incoming sensory information in adverbialist terms.

¹⁶ See Aydede & Fulkerson (2014) for the distinction between the Object and Experience Views that locate the affective qualities differently. In their (2013), they defend a Lockean dispositional account of affective qualities as attaching to the worldly objects of our experiences.

Painful Pains

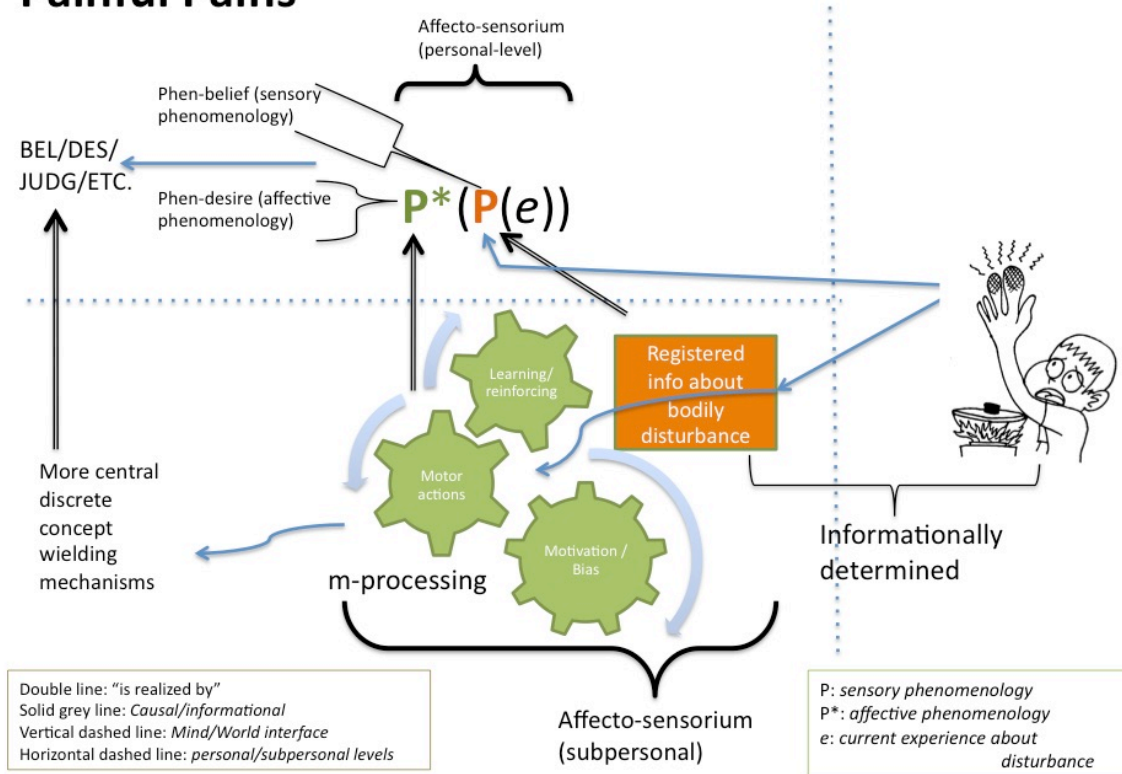


FIGURE: Painful pains according to psychofunctionalism — this basic framework can be generalized to all sensory affect, positive or negative.

All of this is moot, however, if it turns out that psychofunctionalism is also the only game in town. In what follows, we argue that, when modified appropriately the two competing alternative accounts are not, in fact, competitors with the psychofunctionalist view. Instead, they are either gussied up versions of the view when they work, or else they are just inadequate or offer at best metaphors instead of a theory.

How to adjudicate between these three reductive theories? There are basically two ways of doing this. One is to compare the theories against well-known, or at least agreed upon, facts holding of the target domain these theories are theories of. These may be observational, theoretical or normative facts. We can try to find out which theories do a better job in explaining or accommodating these facts. The other method is to describe difficulties internal to each theory. These may relate to the assumptions they make, or their observable consequences, or their formal or theoretical commitments, etc. We can then find out that some of these difficulties make the theories implausible or unlikely, and the attempt to fix them may collapse a theory into another, etc. We have arguments of both kinds against evaluativism and imperativism and for psychofunctionalism. In this paper,

however, we will be concerned only with comparative arguments of the first kind, comparing how they fare about the motivating and justificatory reason providing force of sensory affect.

4 Sensory Affect and Reasons

To simplify the discussion and allow us to focus on the relevant contrasts, let's work on an example. Suppose Sally was making a grilled cheese sandwich and burned her finger accidentally on the hot pan. Let d be the burnt condition of her finger. Let e be the somatosensory experience caused by this in the standard way. Thus, e is Sally's pain experience of her burnt finger and it is deeply unpleasant (painful). According to the three theories then, and simplifying to bring out what is essential:¹⁷

Imperativism: e 's unpleasantness $\approx e$'s incorporating an experiential command to Sally, <see to it that d is no more!>

Evaluativism: e 's unpleasantness $\approx e$'s experientially representing d to Sally as bad.

Psychofunctionalism: e 's unpleasantness \approx m-processing of the sensory information (about d) contained in e . (*Roughly:* phen-believing that d is occurring and phen-desiring that d is no more — see below.)

One way to evaluate these competing theories is to look at how well they explain certain truths about painful pains (and other affective experiences). Two of these (used by Bain himself against the imperativist and attitudinal views) are of special importance precisely because they are thought to pose particularly hard problems for non-representationalist theories. These are the motivational (rationalizing) and justificatory roles of painful episodes (what Bain 2013 calls their “hedomotive” role). Pains not only motivate but also normatively justify certain kinds of behavior as well as the formation of certain beliefs and desires. In other words, they provide motivating and justifying (good) *reasons* — albeit defeasible reasons — for intentional action. Let us explain.

¹⁷ Note that none of these theories are meant to solve (all by themselves anyway) the general problem of the explanatory gap said to exist between phenomenal consciousness and the physical processes that seem to support/realize this consciousness. So we will leave this issue aside — see Aydede & Güzeldere (2005) for a detailed proposal about how to address this general problem.

Another potential worry is that psychofunctionalism may not be in competition with the other representationalist views, since, one might argue, any naturalistic representationalism is committed to some psychofunctionalist realization or other — it all depends on what actual psychofunctional role (more complete and precise specification of CR) is being proposed as doing the explanatory work. We set this worry aside too. This is partly because we strongly suspect that any further elaboration of this functional role along the way CR suggests above won't be suitable as a reduction base for the representationalist proposals expressed by the right hand side of these identities. Another reason is that we suspect that any proposal about how to naturalize representational content will have to appeal to more than (narrow) functional/causal roles. But this is a huge topic we cannot take up here.

Consider first pain’s motivating role. When Sally burns her finger she doesn’t simply experience the pain, *e*. In addition, and *because of e*, she engages in complex intentional actions and forms new intentional states downstream. When we ask what motivated Sally’s actions, we want to know *why* Sally performed all these complex actions. Consider what actions Sally does after she burns her finger: she quickly moves her finger away from the pan, she screams and immediately seeks out ice and cold water for her hand, a bit later she puts soothing and anti-inflammatory cream on her finger. Note that these are pieces of behavior directed toward the condition of her finger, *d*. Call this complex set of behavior *d-directed* behavior.¹⁸ But this is not all that happens: her intentional mental states also undergo complex changes *because of e*: her attention is drawn immediately to the injury, the experience generates stress and concern, and produces in her a set of beliefs and desires.¹⁹

- (d-Belief_i) a belief that she burned her finger (i.e., *d* occurred),²⁰ and
- (d-Belief_n) that this is bad for her (i.e., *d* is bad for her), and
- (d-Desire) a desire to get rid of *d* (i.e., that *d* cease or be no more).

Sally also has strong experientially directed beliefs and desires:

- (e-Desire) the desire that the pain experience, *e*, itself stop, and
- (e-Belief_n) a belief that *e* is bad for her,

and as a result, she also engages in behavior directed to stopping the pain itself — for instance by taking pain killers — call this *e-directed* behavior.

Sally’s pain experience not only *causes* her *d*- and *e*-directed behavior but also *rationalizes* them in the sense that it makes them intelligible and rational from her subjective perspective. Given her experience, it makes rational sense to engage in these kinds of behavior. So when we say that *e* motivates Sally’s behavior, we don’t have mere

¹⁸ Depending on the context, *d*-directed behavior can be protective, rehabilitating, comforting, tending, terminating, stopping, etc. We will sometimes use ‘*d*-avoidance behavior’ for this constellation of behavior patterns.

¹⁹ These are ordinary beliefs and desires with conceptual content. For convenience, we use ‘d-’ and ‘e-’ constructions to indicate that these beliefs and desires are directed towards the disturbance, *d*, and the experience, *e*, respectively. We use subscripts ‘i’ and ‘n’ as gnomonic devices to mark that the belief is indicative (descriptive) or normative (evaluative), respectively.

²⁰ And the belief that this is *why* she is feeling a *pain in her finger*. This sort of belief attributing pains to body parts is particularly problematic for representationalists. But we will leave this issue aside in this paper (see Aydede 2009a).

causation in mind but also rationalization of action.²¹ Sally's pain experience provides reasons *for which* she acts the way she does.

Furthermore, these are *good* reasons that rightly *justify* Sally's actions of both kinds. When we ask what *justifies* Sally's actions, we are not seeking a merely motivational/rationalizing story from her own perspective. Instead, we want to know if the actions were performed for *good* reasons.²² The actions that Sally performs are justified by her painful experience and her burnt finger. In other words, her painful experience and what it indicates give her perfectly *good* reasons to seek ice, avoid the hot pan, *and* take pain killers.

We have, then, the following claims ordinarily believed to be expressing self-evident truths:

(d-Motivation) *e*'s painfulness *inherently* (but defeasibly) motivates Sally's *d*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.²³

(d-Justification) *e*'s painfulness, insofar as *d* is in fact bad for her, *provides* Sally with a good reason for her *d*-directed behavior — it *justifies* her *d*-directed behavior (as well as her belief that *d* is bad for her d-Belief_n).

(e-Motivation) *e*'s painfulness *inherently* (but defeasibly) motivates Sally's *e*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.

(e-Justification) *e*'s painfulness is *intrinsically* bad, which is to say that it *constitutes* a good non-instrumental (i.e., justifying) reason for both Sally's *e*-directed behavior, and her e-Desire and e-Belief_n.²⁴

²¹ Cf. Bain (2013) and O'Sullivan & Shroer (2012). In explaining intentional behavior, we follow the Davidsonian tradition in assuming that causal and rationalizing explanations don't compete — they are different descriptions of one and the same process (*perhaps* at different levels).

²² That I believe my watch is in the drawer is a *motivating/rationalizing reason* for why I open the drawer. But given that my watch is in fact on the kitchen counter, it doesn't provide me with a *good reason* for my action: it fails to normatively *justify* my action by dint of being false. Also, in what follows, when we use 'motivation,' we will include *rationalization* in its intension.

²³ That is, without the need for any beliefs and desires *additional* to the stock of Sally's beliefs and desires she already had roughly at the time at which she started having the pain. We are not claiming that it's possible to have motivating pains without possessing *any* beliefs and desires — perhaps only intentional systems can be properly *motivated*.

²⁴ This doesn't rule out the possibility of painfulness being instrumentally good, and therefore desired.

It is certainly not out of question to deny these claims as expressing truths. Indeed, Klein (forthcoming) seems tempted to deny both (d-Motivation) and (e-Motivation). Cutter and Tye (2104) flat-out deny both (e-Motivation) and (e-Justification).²⁵ But given a modest phenomenological realism about sensory affect, we will not question these claims. Our concern is to find out which of the three accounts does a better job of explaining why they hold.

4.1 Evaluativism and Reasons

One of the major motivations for representationalist views of affect is the thought that no state can rationalize or justify behavior and the formation of other intentional states without possessing a representational content. So, for instance, a major complaint against felt-quality views of sensory affect is that the occurrence of a mere phenomenal character or a quale cannot inherently rationalize or intrinsically justify — a felt quality is simply not the right kind of thing for that. So the evaluativists think that their account is best suited to explain why the above claims hold. Affect can motivate/rationalize *and* justify because it is the experiential representation of an evaluative state of affairs. If we replace ‘*e*’s unpleasantness’ with ‘*e*’s representing *d* as bad’ in (d-Motivation) through (e-Justification), we obtain the evaluativist formulations of the claims that are supposed to remain true. Let’s start with

(d-Motivation_{eval}) *e*’s representing *d* as bad inherently (but defeasibly) motivates Sally’s *d*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.

This seems to be a version of motivational internalism — the view that representing evaluative/normative state-of-affairs in a truth-apt way has an internal connection to motivation. This is a controversial thesis, but we do not want to rely solely on this for our criticism since we will offer a more direct criticism shortly. We are officially neutral about motivational internalism. For the moment, we simply note that evaluativism seems to require the truth of this controversial thesis (cf. Cohen & Fulkerson, 2014).

What about (d-Justification)?

²⁵ Corns (2013) also seems to deny d-Motivation. But this is because she thinks, following the experimental results of Berridge and his colleagues (see, e.g., Peciña *et al*, 2006), that hedonic value (pure pleasantness) can be separated from the motivating role of certain sensations. However, the personal-level or phenomenological significance of these results is moot. Berridge and his colleagues certainly seem to have unearthed separate subpersonal mechanisms that underlie motivation, reward, and learning. But we have doubts about the phenomenological claims made directly on the basis of these mechanisms. Martínez (forthcoming), after distinguishing motivation from justification as well as damage-directed behaviour from experience-directed behaviour in a way similar to our four claims above, also seems to deny the last three claims — although he pitches his account explicitly in terms of whether *pains* can provide the relevant range of *reasons*.

Sally also has a strong desire

(e-Desire*) that the pain experience, e^* , itself stop, and

(e-Belief_n*) a belief that e^* is bad for her,

Note that g-Desire is now *not* rational (even from her own subjective perspective) unless we provide a different rationalization from Sally's perspective. A rationalizing belief is not difficult to find:

(g-Belief_{n-r}) that g is bad for her (in the sense that g is causing e^* , which *is*, and is believed by Sally *to be*, bad).

We can mark this relational sense of 'bad' (causing painful pains) with 'r-bad'. Note that the sense in which painful pain experiences are themselves bad, as e-Belief_n and e-Belief_n* express, is still different. Painful pains are bad not because they are themselves causing painful pains. Presumably they are also not physically bad — at least not in any sense in which their objects, physical disorders, are physically bad. We may capture this third sense by saying that the painful pains or pleasant sensations are *intrinsically* bad or good. We will say more about this below.

But now note the following. This second scenario suggests that the rationalization and justification of Sally's desire to get rid of d (d-Desire) and the subsequent d -avoidance behavior, in the original scenario, has in fact very little to do with the evaluative content of her experience, e , that d is bad in the physical or some objective bodily sense. Because the second scenario suggests that even if e misrepresented d as bad, Sally's desire (d-Desire) and her d -avoidance behavior would still be rational and justified as long as getting rid of d is a way of getting rid of her painful pain.²⁷ This raises the question of which belief is rationalized and justified by Sally's pain experience in the original scenario where she burns her finger? The belief

(d-Belief_n) that d is bad for her (in the physical sense),

or the belief

(d-Belief_{n-r}) that d is r-bad for her
(i.e., d is bad in the sense of causing a bad experience)?

Intuitively the right answer is: both — although d-Belief_n may now have less direct justification (see below). But the evaluativist cannot agree that d-Belief_{n-r} is rationalized and justified by the representational content of e . For the evaluative content of e is not the content of d-Belief_{n-r}. The evaluativist needs to explain why painful pain experiences, i.e., experiences that merely represent a body part to be physically bad, are themselves bad so

²⁷ In this case, though, the justification involved would not be direct and the reason provided by the experience's evaluative content would be instrumental.

that much of our body-directed desires and avoidance behavior would remain the same (rationally and justifiably) even when our painful pains misrepresent. This the evaluator cannot do by merely advertent to the existence of evaluative content. Here is another way to bring this out.

Suppose Sally doesn't have any opinion about whether she has developed allodynia — indeed she doesn't even know what allodynia is. For all she knows gentle touch may be causing some physical damage (given the slight swallowing and reddening of the over-sensitive area), or maybe not. Still, when slightly touched, she feels pain, and forms the belief that

(B) *that* is bad,

where 'that' refers to the gentle touching, *g*. We assume that Sally knows what she means and is intuitively justified in her belief by her experience. But Sally is surely *not* prepared to withdraw her judgment (B) if she learns that *g* is not damaging or bad in any physical or objective bodily sense. Just tell her "but, Sally, gentle touching is not in any physical sense bad, damaging, or disturbing, you are just wrong in thinking *that* is bad for you." The most likely response you will get is "it is bad, it still hurts!" But this just means in her mouth that *g* is causing a painful pain in her. This seems to show that Sally has in mind (g-Belief_{n-r}) from the beginning — in other words, even if, contrary to imagined facts, *g* turned out to be physically bad/damaging.

However, if epistemic justification is (at least, partly) a matter of representational content and evaluativists are right about the content of pain experiences, Sally's experience, even if it may rationalize it, does not justify (B) — so none of her *g*-directed subsequent behavior is justified either — if (B) attributes physical badness/damage to *g*. For (B) is just false according to evaluativism under the imagined scenario (just as the evaluative content of her pain is), despite Sally's protests to the contrary.²⁸

It may be that in normal cases both the experientially represented damage and its experiential representation (pain) are bad. But evaluativism seems incapable of explaining why our beliefs and desires subsequent to our painful pain experiences seem to be aligned in their content with the intrinsic badness of experiences and the r-badness of their causes, rather than the physical badness of the damage. So contrary to their own advertisements, evaluativists do not have an adequate account of how painful pain experiences provide justifying reasons. This brings us to the other two truths of the four stated above. According to evaluativism, the proper readings after the required substitution are:

²⁸ Can the evaluator claim that the evaluative content of painful pain experiences is that *d* is r-bad for Sally? In other words, the unpleasantness of Sally's *e* is just its experientially representing *d* as r-bad. But this would be circular. It just comes to the claim that the unpleasantness of *e* is just representing *d* as causing an unpleasant experience, *e*. (Cf. Bain's own criticism of what he calls the Dislike View in his 2013.)

(e-Motivation_{eval}) *e*'s representing *d* as bad *inherently* (but defeasibly) motivates Sally's *e*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.

(e-Justification_{eval}) *e*'s representing *d* as bad is *intrinsically* bad, which is to say that it *constitutes* a good non-instrumental (i.e., justifying) reason for both Sally's *e*-directed behavior, and her anti-*e* desire (e-Desire) and belief (e-Belief_n).

Far from being self-evident, these do not seem to be true at all.²⁹ Philosophically, one of the most interesting and important things to be explained about sensory affect is why the affective experiences themselves are pleasant or unpleasant, and therefore, intuitively, are non-instrumentally good or bad to have. It seems to us that evaluativism does not touch this issue at all. Focusing on the case of pain sometimes obscures this point since normally painful pains signal the objectively bad things happening to parts of the body. Because of this, one's intuitions do not usually settle the claim that if it weren't for the badness (unpleasantness) of pains, what they indicate wouldn't be bad or unpleasant. Consider the intuition that if it weren't for the goodness or pleasantness of pleasant sensations, the objects of these sensations (e.g., the strawberry or its taste objectively understood) wouldn't be pleasant or good. It is clear that the taste of a strawberry (objectively understood), or the strawberry for that matter, is good or pleasant only in the sense that it is pleasure-causing, i.e., what is good or pleasant in the first instance is the experience of pleasure — the pleasantness of the taste sensation. Similarly, in the pain case: what is important from the point of view of understanding our pain-based beliefs and actions is the badness of painful pains, not the badness — or represented badness — of the damage or other stimulus.³⁰ It seems to us that the intuitions about the *primacy* of affect as attaching to the experiences/sensations are fairly secure (see our 2014 for more elaboration). Saying that this experiential affect is merely the representation of their objects as good (objectively or physically), or as pleasant for that matter, does not take us far.³¹

David Bain (2013) briefly responds to the worry about how the possession of *d*-directed evaluative content can make the experience itself intrinsically bad and can therefore motivate *e*-directed behavior. As we argued above, this is particularly worrisome when the evaluative content is false and known to be false by the sufferer. As Bain puts it:

²⁹ Again, if an evaluativist holds a relevantly modified version of motivational internalism, *perhaps* (e-Motivation_{eval}) may be retained without too much trouble. But motivational internalism is controversial, and we are not sure that evaluativists would be happy to rest the plausibility of their case on its truth.

³⁰ This is, of course, *not* to deny that the physical disturbances are important and usually objectively bad for the organisms who register them.

³¹ In fact, it does not take us anywhere. See Aydede & Fulkerson (2014) for elaboration of this criticism of representationalism about affect.

“Why shoot the messenger if you know the message is false?” (2013: S86).³² He points out the existence of other instances where it seems very natural to think of these states being themselves bad for the subject. He claims that we have an intuitive understanding of why, for instance, being in emotional states such as grief or fear is itself intrinsically bad for the experiencer. These are cases in which something *experientially* strikes the subject as bad, and because of this, Bain seems to claim, being in such states is itself intrinsically bad for the subject.

But even if we accept the analogy with emotions (which is moot), it only shows that there are other cases where we may be paying more attention to the messenger than to the message. In fact, we don’t need to refer to emotions for good examples for this. Pretty much any experience with medium to strong negative or positive affective character will do. When I find the taste of strawberry so pleasant, or am having an orgasm, I am very much into the messenger, not (merely) the message (i.e., these are states I have direct interest in because they themselves are intrinsically good or bad to be in). But this is just to repeat the main worry: how can experientially representing bodily badness or goodness itself be intrinsically good or bad? But we thought that an essential part of the philosophical project about sensory affect was to explain just this on evaluativism: how could representing bodily badness or goodness metaphysically amount to experiential unpleasantness (badness) or pleasantness (goodness) and can therefore motivate and justify our *e*-directed behavior and propositional attitudes (e-Motivation and e-Justification above)? Pointing to even more problematic examples like positive and negative emotions as existence proofs amplifies the mystery rather than answering it.^{33,34}

³² See also Jacobson (2013) for a nice elaboration of this criticism. Cf. Aydede (2005, 2009b). Cutter & Tye (2014) respond to Jacobson by denying (e-Motivation) and (e-Justification).

³³ Is it plausible to reject (e-Justification) — or (e-Motivation) for that matter — on the ground that the badness of pain experiences is completely extrinsic — for instance, by saying that it derives it from the badness of its consequences such as being exhausting or distracting? (Cf. Cutter & Tye, 2014; Martínez, forthcoming). We cannot argue against its plausibility in this paper, but if evaluativism can be saved only by denying (e-Justification), it’s not worth saving, especially when there are clear and better alternatives to it, such as psychofunctionalism. In general, we believe that our own naturalistic view will be preferable to any other naturalistic view that denies that painful pains are intrinsically bad. We invite the reader to do a comparative assessment and leave the matter at that.

³⁴ In a forthcoming article, in responding to an anonymous reviewer, David Bain seems to suggest that his evaluativist view can be formulated also as an experiential desire view where the desire is directed towards bodily disturbance, *d*. We note here that these views are not equivalent if the represented badness of *d* is physical or objective badness not constituted entirely by *r*-badness. If, on the other hand, the represented badness is *r*-badness, then evaluativism is circular, but there is no corresponding difficulty of an experiential desire view — indeed it is not at all clear what an experiential desire has to do, *content-wise*, with the *physical* badness of a disturbance like *d* or *g*. Also, and relatedly, if these views were equivalent, there would be no ground for worrying about providing a psychosemantics for the evaluativist content. The experiential desire view, on the other hand, does not need a psychosemantics beyond the psychosemantics for the indicative content that *this* disturbance is occurring: being aversive to this content is a *functional* affair that does not need a psychosemantic explanation. Similarly, there would be no ground for worrying about the messenger-shooting problem: the experiential desire view delivers a quite natural and plausible account of why the experience itself is bad in a way evaluativism does not. But Bain, quite rightly, worries about these two

4.2 Advantages of Psychofunctionalism

Note that psychofunctionalism does not suffer from *any* of these problems. According to our view, negative or positive affect is the m-processing of incoming sensory information, which is, roughly, the phen-desiring or phen-aversion of what the sensory information is about. Because the incoming information is encoded in sensory code being subject to m-processing (with something like CR), pleasant sensory experiences are pleasant because their pleasantness is constituted, *very roughly*, by the satisfaction of phen-desires as registered by the simultaneous phen-beliefs. Unpleasant sensory experiences, including painful pains, are unpleasant because their unpleasantness is constituted, *very roughly*, by the frustration of phen-desires as registered by the relevant phen-beliefs.³⁵

Even given this much, we can explain how the painfulness of Sally’s pain, *e*, motivates/rationalizes as well as justifies both Sally’s *d*-directed and *e*-directed propositional attitudes as well as behavior. Her *d*-directed propositional attitudes,

- (d-Belief_i) her belief that she burned her finger (i.e., *d* occurred), and
- (d-Belief_n) that this is bad for her (i.e., *d* is bad for her), and
- (d-Desire) her desire to get rid of *d* (i.e., that *d* cease or be no more),

are rationalized and justified by her experience, no matter how ‘bad’ is interpreted — although we usually have relational badness in mind. Sally’s experience presents *d* (or *g* for that matter) as the object of her frustrated phen-desires. Similarly, Sally’s *e*-directed propositional attitudes,

- (e-Desire) the desire that the pain experience, *e*, itself stop, and

problems deeply. (In personal correspondence, Bain has kindly pointed out that he is officially neutral about whether desires can be analyzed as evaluative beliefs of a certain sort and that our phen-desire view is indeed not equivalent to his evaluative content account of pain’s affect.)

³⁵ The qualification “very roughly” is required because the description in terms of phen-desires and phen-beliefs are close to personal level descriptions and may not neatly correspond to any natural delineation of mechanisms constituting m-processing. For instance, what is phen-believed may be that a certain taste objectively understood is present in my mouth as a result of, say, biting a ripe strawberry (descriptive sensory phenomenology). But this content is *also* what is phen-desired (affective/conative phenomenology). Are there further mechanisms during m-processing that additionally register the fact that what is phen-believed satisfies what is thus phen-desired — a state analogous to a meta-belief? Whether this be so depends on further empirical facts about how the affective system works as a network between informational, motivational (drive), and learning mechanisms. We won’t further explore the details here. Suffice it to say that the phen-belief and the phen-desire involved in a given episode of sensory affect are functionally defined over the same sensory signal carrying information about, more or less, what is happening at the location of the relevant sensory receptors. Without a common binding and the availability of this information, it is hard to fathom what the point of having an affective system may come to. We do suspect, however, that the consciously registered (dis)pleasure may arise out of an engineering need to have an explicit signal for desire (dis)satisfaction: an autonomous system capable of individual reinforcement learning needs to know when and how its desires or phen-desires are satisfied or frustrated as a consequence of its own actions.

(e-Belief_n) the belief that *e* is bad for her,

are rationalized and justified by her experience, although in a different way. How?

Recall what we said earlier: if there is a desire involved in affective experiences, it is a phenomenologically salient *experiential desire*. We called this a *phen-desire*, and contrasted it with the sensory informational content of a perceptual experience (called *phen-believing*). On this model, the affective phenomenology is one primarily of phen-desiring and its continuing frustration (or satisfaction) as registered by the incoming sensory information (phen-believing); the primary locus of the relevant sensory affect is the experience itself. So it follows that Sally, in having a phen-desire, is undergoing (and aware of) an experience that is intrinsically bad (the negative m-processing of sensory input — roughly: phen-desire frustration — is itself bad). Her forming e-Desire and e-Belief_n (as well as her introspective belief that *e* is occurring) shows Sally's direct (intuitive) cognitive and conative grasp of this normative fact.³⁶

So when Sally experiences the sharp burning pain, she is completely rational in seeking to end that experience. This is so even if there is no damage at all in her body, or any threat of damage, or the like. The experience itself is negative (unpleasant, intrinsically bad), and one she is both rational and justified in ending.

More importantly, her phen-beliefs and phen-desires also motivate and justify her d-directed behaviors. Sally's phen-belief (the descriptive sensory phenomenal content) that *d* is occurring rationalizes, and if correct, justifies d-Belief_i. When she burns her finger she becomes experientially aware that *d* is happening, and so she is rational to act in ways reflecting that awareness. d-Belief_n is motivated and justified in virtue of the fact that her experience is constituted by both the phen-belief concerning the disturbance and a frustrated phen-desire about this disturbance. Recall that *d* is simultaneously the object of her phen-desire and phen-belief. (It is the very same sensory signal that is subjected simultaneously to both indicative and conative processes, so to speak.) While the negative phenomenology primarily attaches to the experience constituted by her frustrated phen-desire, her phen-desire itself is directed at the disturbance *d*. So, her phen-desire (the affective phenomenal content) that *d* be no more both rationalizes and justifies her d-Desire. Note that Sally's phen-desire and d-Desire occur at different levels (the former is experiential, the latter is a propositional attitude), and thus do not compete with one another as either causes or reasons. It is perfectly natural to say that Sally has her d-Desire *because* of her phen-desire, and that the phen-desire *provides a good reason* for her d-Desire. It is also perfectly natural to say that Sally's frustrated phen-desire *constitutes a good reason* for

³⁶ We are thus in agreement with Goldstein (1980, 1987) that the intrinsic badness of pains and the intrinsic goodness of pleasures are manifest to rational animals (who, we would like to add, have the requisite conceptual resources).

her e-Belief_n and e-Desire. There is no threat of circularity or emptiness in either of these formulations.³⁷

We now have fairly straightforward experiential equivalents of standard belief-desire explanations (both causal and rationalizing), and, as we should expect, the normative justification goes along with the truth-value of their content. Clearly, if we are right about the rationalization and justification of Sally's relevant propositional attitudes, how to extend this account to cover how both *e*-directed and *d*-directed *behavior* are motivated and justified should be clear. Furthermore, our account has the additional advantage of showing how the relevant range of behavior can be motivated even in the absence of the *actual* formation of propositional attitudes both about *d* and about *e* — if that is possible.³⁸

Finally, a few words about the metaphysics of intrinsic goodness or badness. We identified sensory affect, pleasantness and unpleasantness, with the satisfaction and frustration of phen-desires respectively, which are in turn identified with positive or negative m-processing of incoming sensory information. One may reasonably ask: what is it about the frustration of a phen-desire that makes it intrinsically bad? Granted, painful pain experiences are intrinsically bad — maybe obviously so. But it is not obvious that phen-desire frustration that we claim makes pains painful (unpleasant) is intrinsically bad. This is a legitimate worry for those inclined towards non-reductionist or non-naturalist accounts of any sort of intrinsic value. Sensory affect, as we understand it, is a species of intrinsic value — we are officially neutral about whether there are other species and whether they can be reduced or naturalized if there are.³⁹ But we do want to claim that the

³⁷ A further substantial advantage of this view is that it may offer a plausible account of sadism and other cases of disconnect between pain and our all-things-considered desires.

³⁸ Additionally, our psychofunctionalism seems to provide a potential resolution of certain core aspects of the debate between (normative) reason internalism and reason externalism. It is hard to say without generating controversy what these views are and we will not attempt it here. Suffice it to say that the latter view, as we understand it, implies that all desires admit of independent justification, while the former denies this. So, suppose we have naturalistically secured the notion of experiential desire and belief — as indicative and conative/desiderative representation in sensory code (as sensory information being m-processed). Then just as sensory/perceptual experiences can justify our perceptual beliefs, experiential desires (phen-desires) can rationalize and justify our desires as conceptually structured propositional attitudes (call these, c-desires). Reason externalism could be defended only as applied to c-desires — all c-desires do or can admit non-arbitrary independent reasons. For instance, the reason for c-desiring to get rid of pains is that pains are intrinsically bad (continually frustrated phen-desires) and we can grasp them as such (cf. Goldstein 1980). But then there wouldn't be any reason internalism as applied to phen-desires — token phen-desires just aren't the kind of states that can have reasons unless they are perceptions of objective sensuous value (à la Johnston, 2001) in which case they would become phen-beliefs about objective sensuous values — but we set this view aside here. Reason externalism would thus win the debate but the insights of reason internalists would nevertheless be preserved — again, assuming that sensory desiderative phenomenology is not the registration of objective value beyond one's skin. Importantly, we can also objectively explain the existence of phen-desires themselves (their *raison d'être*, in other words) on the basis of how or why evolutionary selective pressures worked out the way they did. This may amount to their “justification” as types — but this sort of justification is not at issue in the debate between these two parties.

³⁹ Furthermore, the intrinsic value we have in mind comes in degrees and doesn't compete with the sensory affect's also having instrumental or extrinsic value. In fact, for just the dialectical purposes of this essay, we

intrinsic badness or goodness of a sensation (qua sensation) metaphysically consists of phen-desire frustration or satisfaction as registered by the relevant simultaneous phen-beliefs. The intuition here is that *if* sensory affect's having intrinsic value is not an ontologically primitive occurrence, *then* our proposal is the best reductive account available for its metaphysical nature. The possession of intrinsic value by sensory affect, we claim, is made more intelligible by our reductive account. We cannot defend this claim here. But, as naturalists, we find it intuitively plausible and illuminating.

4.3 Imperativism

The imperative view holds that *e*'s painfulness is reduced to the fact that S receives a command from her relevant sensory system to the effect that S engage (or, stop engaging) in a certain sort of behavior directed to the bodily location where the pain is felt. Imperativists differ on whether pains are entirely constituted by commands (Klein) or also contain sensory information (Hall and Martínez). For our purposes nothing much hangs on this issue, since on *both* accounts it is supposed to be the commands that explain the motivational and rationalizing role of pains. The sensory content is at best value neutral. So on all imperative accounts it is the *command* that is supposed to give a reason for our behavior.

We have been assuming that it makes sense to talk about experiential representations with the general understanding that the phenomenology of perceptual/affective experiences can be reductively explained by appeal to such representations and the way they are processed. Imperativism appeals to experiential *commands* — experiential representations in imperative mood. We will assume that this view is different from views that appeal to experiential *desires* or causal/functional roles — e.g., phen-desires — such as ours where it is the psychofunctional role (m-processing \approx CR, above) of incoming sensory information that turns the resulting mental process into a desire-like process directed toward the object of the experience about which the information is gathered through sensory means. The difference is, roughly, analogous to the difference between the following two scenarios.

The police shout at the protesters: “Disperse and leave the square immediately!” This is a serious command issued by a (usually) legitimate authority. This command may or may not be obeyed by the protesters. But whether it is obeyed is not a direct causal consequence of the command. The command is just one element in the network of causes that result in the actual dispersion of the crowd — if they decide to disperse.

But now suppose the protesters don't comply, perhaps because they fundamentally disagree with the rational ground of the commands or they want to make their voices be heard more strongly for their cause, or whatever. So the police start using powerful water

may even restrict the intrinsic value to just non-derivative or final value *for the agent* whose experience has the affect.

cannons on the protesters and fire tear gas at them. Despite disagreeing with the rational grounds for the police action, the protesters now have a perfectly good reason to leave the square! In fact, in a certain sense, they have no choice in the matter: we can imagine that the water cannons are so powerful that they are almost sufficient to physically disperse the crowd.

Clearly, the police action (shooting high-velocity water streams in big volumes at the protesters) is no longer a command (except perhaps *metaphorically*). Rather, it is an act to *force* the crowd to disperse. As such, once executed, it may fail or succeed in *causing* the crowd to disperse. The police action has an immediate causal/functional efficacy that the command lacks.

We consider the imperativist view of issuing commands in sensory code to be *analogous* to the police command and the experiential desires to the police action. The latter is essentially a matter of causal/functional role. The former is not.⁴⁰ When you look at the candidate description of CR we tentatively gave above to illustrate the kind of psychofunctional role the incoming sensory information must be subjected to in order to constitute affect, you will see that m-processing, although may not often result in actual action, is nevertheless inherently causally efficacious in affecting and biasing all sorts of learning, motor, and decision-theoretic mechanisms — it *changes* processes in these mechanisms and the way they interface with behavior.

It is possible that the defenders of imperativism are open to interpreting their view as compatible or indeed requiring some form of psychofunctionalism. Indeed, Hall (2008) as quoted above seems quite friendly to the idea of cashing out the notion of a command functionally. Similarly, Klein writes: “An animal who has not had food or drink in days does not need to be informed about her body. She needs to be driven to eat and drink; a biologically effective system for doing so should not bother with the niceties of explaining why. So, too, with pain” (2007: 526). This notion of being “driven” and of coercing action without an intermediary explanation seems thoroughly functionalist rather than representationalist. We think that the *motivational oomph*⁴¹ thus understood cannot be captured by the notion of a mere command to the experiencer. However, if the notion of command assumed by imperativism reduces to a certain kind of causal-functional role, then it is at best misleading to advertise the view as a form of (strong) representationalism — it is at best a form of weak intentionalism where the notion of sensory information plays an important role in the ultimately functionalist explanation of sensory affect, and the talk about commands or imperatives are at best metaphorical — just useful aids for

⁴⁰ However, the imperativist ought to give an account of the imperative mood: what makes a certain representation a command rather than an indicative? We think this can't be done without adverting to the functional/conceptual role (use) of the representational vehicle. It may be that once this role is spelled out fully by the imperativist, the view will collapse into a functionalism about affect. But here we leave this issue aside.

⁴¹ Thanks to Jennifer Corns (2013) for this wonderfully useful and insightful phrase!

understanding the psychofunctional role the sensory information plays to count as affective. The all important notion of intentional *content* does not do much then — the motivational punch is delivered by the functional role, and the view collapses to a form of (psycho)functionalism. In our evaluation, we will thus assume a version of imperativism that is not functionalist but thoroughgoing representationalist.⁴²

With these initial clarifications in mind, let's have a look at the explanations offered by imperativists. With the appropriate substitutions, two of the four claims that need explaining become:

(d-Motivation_{imp}) *e*'s incorporating the command <see to it that *d* is no more!> to Sally *inherently* (but defeasibly) motivates Sally's *d*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.

(d-Justification_{imp}) *e*'s incorporating the command <see to it that *d* is no more!> to Sally, insofar as *d* is in fact bad for her, *provides* Sally *with* a good reason for her *d*-directed behavior — it *justifies* her *d*-directed behavior (as well as her belief that *d* is bad for her, d-Belief_n).

Let's start with a few observations. First, note that the commands almost never incorporate the *rationale* for their own issuance. The rational compliance is almost always a matter of an existing setup or understanding of background conditions and knowledge of the social, cultural, and physical context. Secondly and consequently, the reception of ordinary commands rarely (if ever) motivates and rationalizes actions all by itself. A response to a command by whoever receives it is almost always a complex function of the receiver's background beliefs about the authority, legitimacy, or reliability of the issuer of the command combined with the receiver's background concerns, her perception of the context as well as her standing and occurrent desires and preferences (and her mood). It is this set of background beliefs and concerns that determines whether the reception of command will succeed *motivating* as well as *justifying* the receiver's actions (and the subsequent changes in her beliefs and desires). If this is true and applicable to commands issued in sensory code, (d-Motivation_{imp}) and (d-Justification_{imp}) are just false — imperativism cannot explain (d-Motivation) and (d-Justification).⁴³

⁴² Sometimes, imperativists seem to present their view as offering a different notion of content, *imperative content*, rather than a command. The point of the emphasis seems to be that the explanatory work is done solely in terms of content — it is just that there is more than one notion of content: indicative, imperative, interrogative, etc. These are all contents. Compare this to saying that there is one kind of content (propositions) but that this content can be put in different uses by the different ways their representational vehicles function. There are larger issues about which way is a better way of proceeding from a theoretical/explanatory point of view. Clearly we think the latter way is a better way and find the former somewhat mysterious. But we cannot discuss these larger issues here.

⁴³ Klein (forthcoming) seems to reject (d-Motivation), and thus, (d-Motivation_{imp}). Following Helm (2002), Klein agrees that commands all by themselves don't motivate in the absence of the subject's care and concern

Indeed, remember the propositional attitudes Sally forms upon burning her finger:

(d-Belief_i) the belief that she burned her finger (i.e., *d* occurred), and

(d-Belief_n) that this is bad for her (i.e., *d* is bad for her), and

(d-Desire) the desire to get rid of *d* (i.e., that *d* cease or be no more).

In evaluating evaluativism, we have assumed that *e*'s unpleasantness is both the etiological and justificatory ground of d-Belief_n. Leaving aside how 'bad' is to be understood, this is an intuitively plausible assumption: the experience seems to present *d* as being bad. The intuitive plausibility of this assumption under imperativist interpretation becomes moot, however: it is quite unclear what the reception of a mere command to the effect that *d* cease has anything to do with Sally's belief that *d* is bad for her, especially in the second imagined scenario when Sally forms a belief about *g* — the gentle touch when she is suffering from allodynia:

(B) *this* is bad for me (i.e., *g* is bad for me).

We take it that B is true, when believed by her in the imagined context, whether or not she knows *g* is physically harmless. Again, it is not clear how a command in sensory code to the effect that *g* stop could rationalize and justify B — unless combined with a lot of background beliefs and concerns and a heavy reliance on inference, but this would falsify (d-Motivation) and (d-Justification).⁴⁴

about the immediate physical integrity of her body. But such care and concern certainly entail that the subject has a *desire* to protect her body. It seems likely that Klein thinks that this desire is hard-wired or is a standing or dispositional desire of some sort. The trouble with this suggestion is that if the subject has a desire of this sort, all she needs is the information that her body is being physically threatened, and this information is already (experientially) supplied by the sensory aspect of the pain. Commands are not needed. (Klein doesn't think that pain experiences have any non-imperative content. But this makes things more difficult for him, not less. So we are putting this option aside.) The proposal then at best comes down to the experiential equivalent of a belief-desire explanation/rationalization of the subject's subsequent *d*-directed behavior — there is no mystery about how such states can motivate and rationalize, and there is no explanatory need for commands or imperative content.

⁴⁴ On Klein's austere version of imperativism, *e* doesn't incorporate any direct sensory-discriminative (indicative) representational content about the condition of the body. So on this view, a belief like (d-Belief_i) — and probably the desire (d-Desire) — is not *directly* justified by the content of the experience. Experience (i.e., the command in sensory code) plays only an indirect justificatory role when combined with background assumptions and preferences in the way in which a soldier can come to inferentially know that the hostilities started upon receiving his marching orders (see Klein 2007: 523). Furthermore, on Klein's view commands aren't really about the (actually or potentially) damaged parts of the body, but rather they are about *how not to use* the body parts involving the damage. This further distances the pain experience from the damage or the nociceptive physical stimulus. On Klein's view pain experiences have very little to do with what is harming or seems to be harming the tissue where the pain is felt. It appears that the sufferer's knowledge of where the pain is felt is also inferential depending on what body parts are commanded not to be used in a certain way and the background knowledge of the sufferer. We find the austere version of imperativism to be implausible, and for that reason, have been assuming a version of imperativism (along with Hall and Martínez) that does not deny the sensory-discriminatory content of affective sensory experiences. Indeed, we are not sure, it even

Similarly, the rationality and warrantedness of Sally's original desire (d-Desire) as well as her second desire,

(d-Desire*) the desire to get rid of *g* (i.e., that *g* cease or be no more)

become at best quite *indirect* and *inferential* depending on what else Sally knows about the physical context, the reliability of the relevant commands, and what else she believes and prefers.

But most importantly, on imperativism, it becomes quite mysterious why Sally forms:

(e-Desire) the desire that the pain experience, *e*, itself stop, and

(e-Belief_n) the belief that *e* is bad for her.

We take it that these are *directly* justified and warranted by Sally's experience, *e*. And so are the following by her experience, *e**, in the allodynia case:

(e-Desire*) the desire that the pain experience, *e**, itself stop, and

(e-Belief_n*) the belief that *e** is bad for her.

But we are in dark about how imperativism can imply or accommodate these truths.

With respect to Sally's *behavior*, things are worse. To be a *good* reason, the content of the command has to align with the goal of our actions. This means that the command has to reference or 'make intelligible' our actions *with respect to d*. As we have been assuming, Martínez's version of the form of the relevant commands does just this: <see to it that *d* is no more!>

Is this a *good reason* for Sally? It does not appear to be, because it is entirely grounded by *d* itself (and/or Sally's epistemic access to it). That is, Sally *already has* a good reason to remove the disturbance: she has a disturbance in her finger and she's aware of it! What more reason does the addition of the command add? Well, none. This is clear from pathological cases: hold fixed the command but remove the disturbance. Does Sally still have a good reason to behave so as to remove *d*? Clearly not. Reverse the case: keep *d* but remove the command. Does Sally have a good reason to remove *d*? Yes. We thus have a double-dissociation and it cannot be the command that supplies the reason for Sally's *d*-directed behavior. The situation is even clearer in Sally's *g*-directed behavior. Note that such dissociations are not rare, and hold no matter what semantic account of imperatives is used to cash out the view.

This, of course, just brings us to what imperativism has to say about the remaining two explananda under the appropriate substitutions:

makes sense to be an austere imperativist about positive affect involved in various gustatory, olfactory, tactile *pleasant sensations*.

(**e-Motivation_{imp}**) *e*'s incorporating the command <see to it that *d* is no more!> to Sally *inherently* (but defeasibly) motivates Sally's *e*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.

(**e-Justification_{imp}**) *e*'s incorporating the command <see to it that *d* is no more!> to Sally is *intrinsically* bad, which is to say that it *constitutes* a good non-instrumental (i.e., justifying) reason for both Sally's *e*-directed behavior, and her anti-*e* desire (e-Desire) and belief (e-Belief_n).

Clearly, the imperativists have even more difficulties with these. To provide a good reason, the content of the command again has to line up with the command. But how is this supposed to work? Must the command to Sally be something like: <see to it that *e* is no more!>? But what is this command she experiences? Well, it's nothing more than *e* itself! So *e* is an experiential command to see to it that that very command is no more. How can such a self-defeating command have any reasons-giving force? It is *like* an officer ordering a cadet: Disobey this command!⁴⁵

What can the imperativist say here? Martínez (2013 SSPP) suggests that “it’s not like pain is whispering its command in my ear, and I get to judge its authority, or its urgency. Nothing prevents pain from simply changing its subject’s TDL from the inside, so to say.” The TDL, on his view, is a to-do list of all the items one is supposed to do.⁴⁶ A command is then cashed out as something that adds an item to the TDL. But what should adding an item to my TDL give me a reason for anything? Suppose Sam assiduously keeps a to do list on his desk. When he’s not looking, a grad student adds <Give Jones an A!> to his list. Does its presence there give Sam a reason to give Jones an A? It certainly doesn’t seem so. Martínez suggests a possibility, “The reason [why it is rational to try to stop the pain when there is no physical threat] is clear: the pain imperative is clogging our TDL, forcing us to attend urgently to seeing to it that a certain bodily disturbance is no more. If we know on other grounds that we do *not* have to see to it, it’s entirely rational and justified to try to prevent the pain from messing with our to-do list.”

The nature of Martínez’s suggestion is analogous to the difference between the police action and the police command we introduced above. Commands themselves have

⁴⁵ Or, as pointed out by David Bain in correspondence, more accurately like <stop this command!>. Imperativists, however, have never invoked experience-directed commands.

⁴⁶ Martínez in his (forthcoming) has dropped the analogy with “to-do-list,” but the basic idea behind it is still in force. In his forthcoming paper, which came to our attention considerably after we’d delivered the present criticism at the 2014 Pacific APA, Martínez explicitly addresses the question of how pains can give or be reasons for behavior. But, as far as we can see, our main criticisms above remain. Furthermore, Martínez seems to deny (d-Justification), (e-Motivation), and (e-Justification). We say “seems” because he is relatively clear only about denying (e-Justification). But we fail to see how he escapes denying the other two, given what he says about what the painfulness of pains comes to and how it motivates or justifies.

limited reasons-giving force unless wedded to some downstream story (clogging our inbox). If the nociceptive information processing is directly connected to the ordering of one's TDL in a way that cannot be more or less rationally/cognitively controlled, it is no longer a command. To go this route is to abandon imperativism, and with it, representationalism. But even worse, this suggestion comes down to admitting that a psychofunctionalist account is the right explanatory account for the unpleasantness of pains. In this story, it is not the *commands* themselves that undergird the motivational oomph, but rather the hard-wired causal role of incoming sensory information in *forcing* the priorities to be ordered differently — this is psychofunctionalism. Furthermore, it is not clear to us how literally we should take this talk of “to-do list”. When taken literally, this seems to require that subjects keep a to-do list for themselves keeping track of which commands should be issued to oneself in which order. But why should one have a list of self-commands? All the issues that we raise about commands in ordinary contexts arise here too. What reasons might one have to obey such commands one issues to oneself, etc.? It appears that the talk about TDL is just loose talk about having beliefs, desires, preferences, intentions, decisions, volitions, etc. or whatever the appropriate subpersonal computational decision-theoretic mechanisms and processes are involved.

Finally, Martínez's suggestion comes down to a rejection of (e-Justification). In effect, Martínez tells us that pain experiences aren't *intrinsically bad*, but they are bad for their inconvenient consequences — they, for instance, tend to clog our to-do list. (If, on the other hand, Martínez suggests that this is not a *consequence* of our painful experiences but rather is part of what it *is* to have a painful experience, then again, the suggestion is not imperativist anymore, but psychofunctionalist — and we are sympathetic.)

5 Conclusion

Despite our criticisms, we believe that all three theories of sensory affect present views that are close to each other. All are naturalistic and offer *a posteriori* accounts of sensory affect playing with more or less the same naturalistic resources available to them. We believe that at the core of these resources is the functional/causal role that the incoming sensory information plays. We labeled this complex set of processes defined over sensory information *m-processing* (with a role like CR), and identified it, roughly, with a more or less personal-level experiential desire, *phen-desire*, directed to the physical stimulus responsible for the relevant sensation. The satisfaction or frustration of this phen-desire as registered by the sensory component is meant to capture the positive or negative affective phenomenology of sensory experiences.

We are not anti-representationalists. Although we reject strong representationalism à la Dretske and Tye, we welcome information-theoretic treatment of sensory and cognitive processes. We think that sensations represent worldly properties and magnitudes, but their phenomenology is not exhausted by their representational content. This is particularly true

for affective phenomenology, which, according to us, is not representational at all. The metaphysics of sensory affect is the functional role the sensory information plays in the mental economy of the agent. Of course, experiential desire or aversion towards a stimulus needs to represent the stimulus. This is done by the sensory signal whose personal-level registration is the sensation (phen-belief), but the mechanism responsible for adding an affective dimension to the sensation is the m-processing of that very *same* signal — its causal/functional role, CR. It is not clear to us that there is any serious alternative to this basic framework. Indeed, it is not clear to us that imperativists or evaluativists do or can reject this framework.

Imperativists seem closer to our view in many ways, but they muddle the core issues by talking about the imperative content or commands, and then worry about developing a semantics for these as if the basic philosophical and metaphysical puzzles surrounding sensory affect would be solved if we had a working semantics for imperatives or commands. This may be a useful thing to have, but we don't see how this project can succeed without essential appeal to the functional/casual role of sensory information to make this semantic story work as a metaphysical account of sensory affect as instantiated in functioning agents.

Evaluativists, on the other hand, need to tell us three things. First, what is it for a physical disturbance to be bad in the way that is required for its experiential representation as bad to be accurate or inaccurate? If we are dealing with a notion of badness other than relational badness (r-badness) or intrinsic badness, we do not have any clue about what it might be.⁴⁷ We have been pretending to know what physical/objective/bodily badness comes to, but our criticism above shows that we do not — neither do evaluativists, it seems to us... Secondly, they need to tell us how representing a physical condition as bad, in whatever physical/bodily sense they have in mind, can motivate in the absence of any beliefs and desires (d-Motivation) without relying on motivational internalism. Motivational internalism, in the context of *sensory affect*, is a mysterious thesis: how could experientially representing something as physically bad all by itself have an internal (necessary) connection to motivation? Unless we assume a functionalist framework of the sort we defend, the existence of such an internal connection is mysterious and puzzling. But if, on the other hand, we assume our functionalist framework, evaluativism becomes otiose.⁴⁸ Thirdly, and perhaps, more importantly, how is representing some bodily

⁴⁷ See our (2014) for a criticism of a proposal by Cutter & Tye (2011). Our criticism extends to any proposal that identifies this badness with a completely objective (extra-mental) property — however complex (historical, relational. etc.) it may be.

⁴⁸ Bain (forthcoming) claims that general care about one's body is an existence condition for one to have painful pain experiences, i.e., for one to have pain experiences that represent disturbed bodily conditions as bad. We find this stipulation also mysterious. If one cares about one's body, in whatever sense relevant here, then one has *desires* about not to have physical disturbances in one's body. But if so, we have a fairly standard (and more economical) explanation of the motivational power of painful pain experiences without relying on identifying the painfulness with experiential representation of disturbance as bad. We can just say

disturbance as bad itself bad? This is the notorious problem of messenger-shooting for all representationalists. The problem is particularly difficult if evaluativists have some sort of objective physical/bodily badness in mind.

This problem, in many ways, is at the very core of philosophical puzzles about sensory affect — it is not just a by-product of certain ways of trying to give a naturalistic account for sensory affect. We are puzzled that all representationalists, while trying to bring philosophical illumination in all sorts of ways, seem to leave the main difficulty more or less untouched.⁴⁹

References

- Alston, W. 1968. "Pleasure." In *The encyclopedia of philosophy*, edited by P. Edwards. New York: Collier-Macmillan.
- Armstrong, D. M. 1962. *Bodily Sensations*. London, Routledge and Kegan Paul.
- Armstrong, D. M. (1968). *A Materialist Theory of the Mind*. New York, Humanities Press.
- Aydede, M. 2000. "An Analysis of Pleasure Vis-à-Vis Pain." *Philosophy and Phenomenological Research* 61(3): 537–570.
- Aydede, M. 2001. "Naturalism, introspection, and direct realism about pain." *Consciousness and Emotion* 2(1): 29–73.
- Aydede, M. 2005. "The Main Difficulty with Pain." In *Pain: New Essays on Its Nature and the Methodology of Its Study*, edited by M. Aydede, Cambridge, Massachusetts: MIT Press.
- Aydede, M. 2009a. "Is Feeling Pain the Perception of Something?" *Journal of Philosophy* 106(10): 531–67.

that the painfulness is the frustrated experiential (or, otherwise) desires implied by this general body care. In fact, whatever the nature of the anti-damage desire implied by body care may be, when combined with the sensory information that bodily damage is occurring, we have a fairly straightforward explanation or rationalization of one's anti-damage behavior. It is not at all clear what we gain by insisting that the painfulness is the representation of disturbance as bad and that the body care and anti-damage desires are only existence conditions for such representations.

⁴⁹ Cutter & Tye (2014) propose to "solve" the difficulty by denying there is any difficulty in the first place: they reject (e-Motivation) and (e-Justification) as false.

A precursor of this paper was written for an invited symposium at the Pacific APA meeting in San Diego (2014) titled "Painful Pains, Yummy Tastes, Stinky Smells: Sensory Affect". We thank the other participants, David Bain, Colin Klein, Manolo Martínez, Amy Kind and the audience members, for their very helpful comments and criticisms that have usefully shaped the subsequent revisions of this paper. We also thank David Bain, Michael Brady, and Jennifer Corns for their comments on an earlier version of this paper as well as for the invitation to participate in a series of conferences on pain in Glasgow and to contribute to this volume. Many of the ideas here have been influenced and helped by the ongoing interactions made possible by these meetings and continuing conversation as a result.

- Aydede, M. 2009b. "Pain." In *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2013/entries/pain/>.
- Aydede, M. 2014. "How to Unify Theories of Sensory Pleasure: An Adverbialist Proposal." *Review of Philosophy and Psychology* 5(1): 119–33.
- Aydede, M. Forthcoming. "A Contemporary Account of Sensory Pleasure." Forthcoming in *Pleasure: A History*, edited by Lisa Shapiro, Oxford UP.
- Aydede, M. & Fulkerson, M. 2013. "Affective Qualities." Presentation at the Pacific APA meeting in San Francisco, April 3013.
- Aydede, M., & M. Fulkerson. 2014. "Affect: Representationalists' Headache." *Philosophical Studies* 170: 175-98.
- Aydede, M., & Güzeldere, G. 2005. "Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness." *Noûs* 39(2): 197–255.
- Bain, D. 2013. "What Makes Pains Unpleasant?" *Philosophical Studies* 166(1): 60–89.
- Bain, D. Forthcoming. "Pains That Don't Hurt." Forthcoming in *Australasian Journal of Philosophy*.
- Berridge, K. 2004. "Motivation concepts in behavioral neuroscience." *Physiology & Behavior* 81(2): 179–209.
- Block, N. 1980. "Troubles with functionalism." In *Readings in philosophy of psychology Vol. 1*, edited by Ned Block, 268–305. Cambridge, Massachusetts: Harvard University Press.
- Brady, M. 2013. "Pain and the Euthyphro Dilemma." Paper presented at the Pain Conference, University of Glasgow, June 18–20, 2013.
- Bramble, B. 2011. "The distinctive feeling theory of pleasure." *Philosophical Studies*, 162(2): 201–217.
- Brandt, R. 1979. *A Theory of the Good and the Right*. Oxford: Clarendon Press.
- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Broad, C. D. 1930. *Five Types of Ethical Theory*. London: Routledge and Kegan Paul.
- Byrne, A., & Tye, M. 2006. "Qualia ain't in the head." *Nous* 40(2): 241–255.
- Byrne, A. 2009. "Experience and content." *The Philosophical Quarterly* 59(236), 429–451.
- Chalmers, D. 2006. "Perception and the Fall from Eden." In *Perceptual Experience*, edited by T. S. Gendler and J. Hawthorne, Oxford UP, 49–125.
- Clark, A. 2005. "Painfulness is Not a Quale." In *Pain: New Essays on Its Nature and the Methodology of Its Study*, edited by M. Aydede, 177–198. Cambridge, Massachusetts: MIT Press.

- Corns, J. 2013. "Unpleasantness, Motivational Oomph, and Painfulness." *Mind and Language* 29(2): 238–254.
- Crisp, R. 2006. *Reasons and the Good*. Oxford, UK: Oxford University Press.
- Cohen, J. & Fulkerson, M. 2014. "Affect, Rationalization, and Motivation." *Review of Philosophy and Psychology* 5(1):103-118.
- Cummins, R. 1981. *The Nature of Psychological Explanation*. Cambridge, Massachusetts: MIT Press.
- Cutter, B., & Tye, M. 2011. "Tracking Representationalism and the Painfulness of Pain." *Philosophical Issues* 21(1): 90–109.
- Cutter, B., & Tye, M. 2014. "Pains and Reasons: Why It Is Rational to Kill the Messenger." *The Philosophical Quarterly* 64(256): 423–433.
- Davis, W. A. 1981. "Pleasure and happiness." *Philosophical Studies* 39(3): 305–317.
- Davis, W. A. 1982. "A Causal Theory of Enjoyment." *Mind* 91(362): 240–256.
- Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, Massachusetts: MIT Press.
- Duncker, K. 1941. "On Pleasure, Emotion, and Striving." *Philosophy and Phenomenological Research* 1(4): 391–430.
- Feldman, F. 1997. "On the Intrinsic Value of Pleasures." *Ethics* 107: 448–466.
- Feldman, F. 2004. *Pleasure and the Good Life*. New York: Oxford University Press.
- Hall, R. 1989. "Are Pains Necessarily Unpleasant?" *Philosophy and Phenomenological Research* 49(4): 643–659.
- Hall, R. 2008. "If it itches, scratch!" *Australasian Journal of Philosophy* 86(4): 525–535.
- Harman, G. 1990. "The Intrinsic Quality of Experience." *Philosophical Perspectives* 4: 31–52.
- Heathwood, C. 2007. "The reduction of sensory pleasure to desire." *Philosophical Studies* 133: 23–44.
- Helm, B. 2002. "Felt evaluations: A theory of pleasure and pain." *American Philosophical Quarterly* 39(1): 13–30.
- Goldstein, I. 1980. "Why People Prefer Pleasure to Pain." *Philosophy* 55(213): 349–362.
- Jacobson, H. 2013. "Killing the Messenger: Representationalism and the Painfulness of Pain." *The Philosophical Quarterly* 63(252): 509–519.
- Johnston, M. 2001. "The Authority of Affect." *Philosophy and Phenomenological Research* 63(1): 181–214.
- Kahane, G. 2009. "Pain, Dislike and Experience." *Utilitas* 21(3): 327–336.
- Kagan, S. 1992. "The limits of well-being." *Social Philosophy and Policy* 9: 169–189.
- Klein, C. 2007. "An imperative theory of pain." *Journal of Philosophy* 104(10), 517–532.
- Klein, C. Forthcoming. "What Pain Asymbolia Really Shows." Forthcoming in *Mind*.

- Labukt, I. 2012. "Hedonic Tone and the Heterogeneity of Pleasure." *Utilitas* 24(02): 172–199.
- Martínez, M. 2011. "Imperative content and the painfulness of pain." *Phenomenology and the Cognitive Sciences* 10(1): 67–90.
- Martínez, M. 2013. "The Rationalizing Role of Pains." Presented at the SSPP meeting in Austin, TX, February 2013.
- Martínez, M. Forthcoming. "Pain as Reasons." Forthcoming in *Philosophical Studies*.
- Moore, G. E. 1903/1993. *Principia Ethica*. In T. Baldwin (Ed.). Cambridge: Cambridge University Press.
- O'Sullivan, B., & Schroer, R. 2012. "Painful Reasons: Representationalism as a Theory of Pain." *The Philosophical Quarterly* 62(249): 737–758.
- Panksepp, J. 1998. *Affective Neuroscience*. New York: Oxford University Press.
- Parfit, D. 2011. *On What Matters*. Vol. 1. Oxford University Press.
- Peciña, S., Smith, K. S., & Berridge, K. C. (2006). "Hedonic Hot Spots in the Brain." *The Neuroscientist* 12(6): 500–511.
- Pitcher, G. 1970. "Pain Perception." *Philosophical Review* 79(3): 368–393.
- Robinson, W.S. 2006. "What Is it Like to Like?" *Philosophical Psychology* 19(6): 743–765.
- Rolls, E. T. 2005. *Emotion Explained*. Oxford University Press.
- Schroeder, T. 2004. *Three Faces of Desire*. Oxford University Press.
- Sidgwick, H. 1907/1981. *The Methods of Ethics*. 7th Ed. Macmillan.
- Sprigge, T. L. S. 2000. "Is the esse of intrinsic value percipi?: pleasure, pain and value." *Royal Institute of Philosophy* (Supp. Volume) 47: 119–140.
- Smuts, A. 2011. "The feels good theory of pleasure." *Philosophical Studies* 155(2): 241–265.
- Tye, M. 1995. *Ten Problems of Consciousness: A representational theory of the phenomenal mind*. Cambridge: MIT Press.
- Tye, M. 2005. "Another Look at Representationalism about Pain." In *Pain: New Essays on Its Nature and the Methodology of Its Study*, edited by M. Aydede. Cambridge, Massachusetts: MIT Press.