

Relational Imperativism About Affective Valence

Antti Kauppinen

Forthcoming in *Oxford Studies in the Philosophy of Mind* Vol. 1

Final Draft, March 31, 2020

It's easy enough to imagine a lion chasing a young antelope on the savannah. Given the failure rate of lion attacks, it is potentially a matter of life and death for both. Their hearts racing, the animals strain their muscles to the utmost, sending dust in the sky and tremors underground. It's much harder to imagine what goes through an ungulate mind, but let's try. Insofar as the antelope has conscious experience, it may well be dominated by fear. Of late, it has become common to think that such affective experiences have the feel they do in virtue of having evaluative content, so that what gives fear its negative feel is that it is as if it is telling the antelope "This is bad" or maybe "Being chased by the lion is bad for you!". In this paper, I'll defend an alternative answer, according to which the phenomenologically distinctive content of the fear is something much closer to "Run, motherf*cker, run!!".

Let's start by clarifying what is at issue. It is widely, though not quite universally, agreed that there is something it is like to have conscious experience – that conscious experience has a *phenomenal character*. What I will focus on here is the phenomenal character of *affective experiences* – algedonic states or aspects, emotions, and moods – which can at least typically be characterized as positive or negative. In other words, affective experiences have *valence*: they feel good or bad. The question I'm going to address, then, is the following: in virtue of what do affective experiences have valenced phenomenal character? As the savannah story suggested, the answer I'll defend will appeal to *imperativ*

content, extending views proposed by Colin Klein and Manolo Martinez in the context of debates about pain and unpleasantness.

My argument will proceed as follows. I'll begin by looking at the functional role of valence – what it does and what doesn't do. I contend that the key role of valence is explaining and rationalizing inherent motivation, but not as currency in deliberation. Here, I focus on two competing intentionalist explanations of valence, evaluativism and imperativism. On the face of it, evaluativists, who hold that affects have valence in virtue of representing something as good or bad, have an advantage when it comes to explaining how valenced experiences rationalize action and judgment, but I raise some doubts about that. Imperativists, in turn, have an edge when it comes to explaining how valenced experiences motivate, but extant views face deep challenges when it comes to other aspects of valence.

In Section 2, I turn to developing a new, generalized version of first-order imperativism that I call relational imperativism. Roughly, I argue that valenced experiences have two kinds of interrelated content. Their *indicative* content represents a state of affairs or an object that matters to the subject in the light of a background concern. Their *imperative* content consists of a command that is subjectively authoritative in virtue of linking up with the background concern, and that tells the subject to see to it that the sort of thing represented by the indicative content either will or won't obtain in the future. I argue that this account provides a suitably nuanced picture of how valenced experience motivates, and a plausible characterization of its phenomenal feel.

The final section addresses some important objections that have been raised in the literature against imperativism. The first important challenge concerns the datum that valenced experiences not only motivate but also rationalize. My response appeals to the well-known psychological mechanism of projection or objectification: when we experience being authoritatively commanded to do something about something, it predictably changes the way

things practically appear to us. When the antelope fears the lion, the lion appears to it as something *to-be-fled-from*. I claim that such *affective practical seemings* resulting from imperatival content suffice to make sense of how valenced experiences rationalize. The second key challenge concerns the explanation of *intensity* of affect in terms of commands. Klein and Martinez have presented a sophisticated model that appeals to a kind of inbuilt priority ranking among imperatives. I propose a simpler account of intensity in terms of the level of opportunity cost that imperative contents tell us to bear in order to satisfy them.

Third, I address what I'll label the Shooting the Officer Objection: why should we want to get rid of experiences like pains, if their negative phenomenal character consists of our being told to see to it that some bodily condition goes away? What's so bad about being told to make the future different from the present? I argue that just as parallel objections to evaluativism, this objection amounts to begging the question when it comes to the badness of negative experiences. I also maintain that the unpleasantness of an experience doesn't in itself motivate us to get rid of the experience, which we can see if we consider world-directed emotions like grief. Instead, what motivates and rationalizes, say, taking painkillers is being (contingently) attitudinally displeased with feeling bad.

1. The Debate About Valence: From Evaluativism to Imperativism

1.1 The Concept of Valence

Let us begin by taking a closer look at the motivational role of valence. If I take pleasure in eating ice cream, I tend to be motivated to eat more ice cream, even if prior to tasting it, I had no such desire. If I'm grateful to Sandy for lending me her bike when I needed it, I tend to be motivated, at least to some extent, to help her in turn, even if I had no such motivation before. Note that you could be motivated to return a favour just because you think it's the right thing to do, without feeling grateful – but your feeling of gratitude seems to suffice on its own for

motivation, at least potentially. If I'm tranquil, I tend to be motivated, if that's the right word, to let things be as they are instead of rearranging them somehow, even if I hankered for something different before.

What these reflections suggest is that *valenced phenomenal character is potentially motivating by itself*. I don't claim that this is obvious – there is more to gratitude, for example, than its phenomenal aspect, and some might argue that it is the non-phenomenal aspects that account for motivation. But insofar as introspection can be relied on here, it appears to be precisely the *feeling* of gratitude that guides me towards reciprocating, or the toothache that finally gets me to call the dentist. It would be rash to conclude that valenced affects necessarily motivate – indeed, I'll argue that they don't. But when they do, they do not merely channel pre-existing desires. I'll use the term “inherent motivation” for this phenomenon (the more common term “intrinsic” suggests necessity). Because of its inherently motivating character, it is unsurprising that it is common to identify valence with pleasantness or unpleasantness of an experience – after all, pleasures and pains seem to be inherently motivating, and motivational hedonists like Bentham even hold that they're the *only* inherent motives. But while I think that positive feelings feel good, I don't believe that valence just *is* pleasantness or unpleasantness, because such identification would suggest that anger, say, inherently motivates *in virtue of* its unpleasantness, which I don't think is true. Moreover, it's not conceptually confused to think of valence as consisting of evaluating something as good or bad, even if one does not think that such evaluations are necessarily pleasant or unpleasant. So instead of pleasantness or unpleasantness of experience, I'll take as my primary explanandum inherently motivating phenomenal character.

Valence does not just motivate action, however – it also *rationalizes* it, at least subjectively. The way I would describe it is that when you have a valenced affective experience, it appears to you that you have a reason to act or should act in a certain way, or at

least the experience presents something in a favourable or unfavourable light. If you're grateful to Sandy, it makes at least some sense to you to return the favour. Your toothache doesn't just get you to call the dentist, but also presents the condition of your tooth in an unfavourable light.

It's worth highlighting that I will not treat it as part of the explanandum the surprisingly common view that valence serves as *common currency in deliberation* (e.g. Levy and Glimcher 2012, Carruthers 2018), that is, as something that commensurates the value of different options. Here is how Luca Barlassina illustrates the alleged deliberative role of valence: "To decide whether to go to the movies or to a restaurant, you can imagine either activity and use the elicited valences ('this seems more pleasant than this') to guide your choice." (MS, 2). But it does not do justice to the phenomenology of deliberation to assume that we make our choices on the basis of how we feel about our options, not to mention how much pleasure or displeasure we anticipate getting from them. While such considerations do play a role in *some* decisions (possibly including the kind of leisure choices Barlassina mentions), deliberation also, and centrally, involves beliefs about what is right or good, and is crucially shaped by our prior plans and commitments. If there's a common currency to deliberation, it's *utility*, and that's just because "utility" is a catch-all term for how our preferences rank our options, whatever they're based on.

1.2 Evaluativism and Its Discontents

How, then, can we account for valence? For reasons of space, I'll have to set aside views that hold there's nothing to explain (e.g. Smuts 2011) and externalist views that appeal to attitudes either towards our experiences or the world (e.g. Heathwood 2007, Jacobson 2019). Instead I'll just focus on two competing *intentionalist* views, which link phenomenal character with the *intentional content* of an experience. Intentionalist views are internalist but not brute,

since they offer a further explanation of valenced phenomenology. According to what I'll call *Phenomenal-to-Intentional Reductionism*, experiences have their phenomenal character in virtue of their intentional content, which in turn can be naturalistically explained (e.g. Tye 1995). A minority tradition subscribes to *Intentional-to-Phenomenal Reductionism*, according to which it's instead phenomenal character that explains intentional content of at least some experiences (e.g. Mendelovici 2018). To be neutral, I'll just assume *Phenomenal-Intentional Parallelism*: the phenomenal character of an experience *matches* its intentional content – same intentional content entails same phenomenal character (at least in those respects that matter here), and vice versa.

The most common form of intentionalism is *representationalism*, according to which mental states have the phenomenal content they do in virtue of their representational content. When it comes to affective phenomenal character, it is not plausible that the representational content in question simply concerns how things are. One reason to think so is that affective character can vary independently of such representations – for example, two people might have pain states that represent the same kind of bodily disturbance while nevertheless experiencing pain of different intensity (Pautz 2010). The standard representationalist answer is to say that affective states also have an additional, *evaluative* representational content, such that variations in character and intensity of affect are explained by variations in the type and degree of represented value – for example, one pain feels worse than another because it evaluates my bodily condition as worse (Cutter and Tye 2011). Here are a few representative evaluativist claims:

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 2012)

The valence component of affective states like pain and pleasure is a nonconceptual representation of badness or goodness. (Carruthers 2018)

An emotion has negative valence because it contains a perceptual experience of a negative thick value. (Teroni forthcoming)

Evaluativism is currently the most popular internalist account of valence. Maybe the best case for it appeals to the role of valence in not only motivating but also rationalizing action (Helm 2002, Bain 2013). I'll reconstruct it as follows:

1. The phenomenal character of an experience matches its intentional content.
2. The intentional content of valenced affects (e.g. pain) rationalizes world-directed actions and judgments about value.
3. The intentional content of valenced affects only rationalizes world directed-actions and judgments about value if they have world-directed representational content, which they represent as good or bad (in a particular way).
4. So, valenced affects represent something as good or bad (stronger: affects have valence in virtue of representing something as good or bad).

Premise 2 calls attention to the fact that an affective experience like feeling pain in your ankle not only motivates but also rationalizes an action that changes your ankle-directed behaviour, such as not putting weight on it. What's more, such experiences also seem to rationalize judging that the damage in your ankle is bad for you, much in the same way as perceptual experiences rationalize perceptual beliefs. Premise 3 then claims that experiences can only play these roles if their contents are evaluative. Suppose that your pain presents the damage in your ankle as bad. This experienced badness rationalizes behaviour that reduces it – by your lights, you have a reason to reduce the damage. It is a further question whether these

evaluative contents are conceptual or non-conceptual, where conceptual contents are those linked with linguistic capacities or general cognitive skills (e.g. Evans 1982). The mainstream view is that they are non-conceptual (Tappolet 2016, Carruthers 2018, Mitchell 2018). One key argument for this claim appeals to fineness of grain: just like our colour experience plausibly discriminates between more shades than we have concepts for, our affective experiences discriminate between more ways of being valuable than our concepts do. As Jonathan Mitchell (2018) maintains, someone with a crude aesthetic conceptual repertoire could nevertheless experience different paintings as beautiful in different ways while lacking even the ability for even demonstrative re-identification of the represented property.¹ It is a tricky question just how non-conceptual contents can rationalize beliefs that do have conceptual contents (indeed, Mitchell thinks they don't), but presumably the answer will be of the same sort as in the case of ordinary perception.²

In spite of its popularity, I have serious doubts about evaluativism, however. First, evaluativists rightly emphasize that one need not possess *concepts* of specific kinds of goodness or badness to have valenced experiences, and that possessing such concepts does not suffice to capture the fine grain of valenced experiences. Even individuals who do have such concepts represent something as good or bad in a different, non-conceptual way when they have affective experiences. But what exactly are we talking about when we talk about non-conceptual representation of goodness or badness? What makes a non-conceptual representation one of *goodness*, in some sense? How do we non-conceptually represent one thing as *better* than another? It is hard to find arguments for this in the literature. It won't

¹ Mitchell's main argument for non-conceptualism is based on first-exposure affective experiences, such as one's first encounter with the sublime, which according to him feature evaluations that are not cognitively significant and which the subject herself consequently can't understand.

² Carruthers also suggests that "only states with nonconceptual content are phenomenally conscious" (2018, 664), so insofar as valenced affects are phenomenally conscious, their content must be nonconceptual. I'm agnostic on this point.

suffice to say, for example, that the subject is disposed to form evaluative beliefs in response to valenced experiences, since that wouldn't explain how the valenced experiences themselves represent their objects evaluatively (you might be disposed to think blue things are good without your experience of blueness representing its object as good). Nor does it suffice if the experiences somehow track what is actually good or bad for the subject (which is the line that Cutter and Tye 2011 try out) – an experience might co-vary with what's good without its object appearing *as good to the subject*. My suspicion is that when we spell out what non-conceptual representation of value amounts to, we end up with a view that is *not* fundamentally an evaluativist one (as I'll argue in Section 3).³

Another worry is an old but good Humean one: how is it that (non-conceptual) representations of goodness or badness are inherently motivating in the way that valence is supposed to be? After all, the standard view in philosophical psychology has it that representational states are not inherently motivating – their job is to tell us how things are, not provide the push (or guidance) in one direction or another (Anscombe 1957). Simply stipulating that they're part of an affective state or have non-conceptual content is no argument to the contrary. There are, to be sure, some who hold that *evaluative* representations can directly motivate. But this is a highly contentious claim, and again raises the question of whether what makes representations specifically evaluative is that they contain a non-cognitive component (e.g. Gibbard 2003, Ridge 2014).

1.3 Varieties of Imperativism

³ Julien Deonna and Fabrice Teroni (2015) have recently developed an alternative evaluativist approach, according to which affective valence is accounted for by embodied evaluative *attitudes* towards non-evaluative contents. From my perspective, the key challenge to this type of view is making sense of the rationalizing role of affective experiences, since it is most naturally explained by appeal to their content.

While the considerations I just mentioned are far from decisive, they do give us a good reason to look for an alternative to evaluativism. The leading intentionalist alternative is *imperativism*. To understand it, let's start with the difference between "The door is open" and "Open the door!". The traditional line is that the sentences have the same content but different force attached to them. But recently many have argued that we need to distinguish a difference in *content* here. Here's how Luca Barlassina and Max Hayward (2019) characterize it:

Indicative Content:

- (i) Has the function of carrying information, for example, that p is the case;
- (ii) Has truth conditions: it is true if p is the case and false if not;
- (iii) The audience correctly uptakes it by forming a belief.

Imperative Content:

- (iv) Has the function of directing its addressee to do something, for example, to F;
- (v) Has satisfaction conditions: it is satisfied if and only if the addressee Fs;
- (vi) The addressee correctly uptakes it by forming a motivation. (Barlassina and Hayward 2019, 6)

Various kinds of argument have been offered for the existence of imperative content (rather than just imperatival force attaching to propositional content). For example, Ruth Millikan argues that the proper function of some content-types, 'imperative intentional icons', is to produce certain effects, and that their content is the "last item of the series of things it is supposed to map onto and to produce" (1984, 100). Hector-Neri Castañeda (1975) holds that some utterances express 'practitions' instead of propositions, where practitions are thought contents that link an agent and action in a practical way (see also Chrisman 2016). Peter Hanks (2007) argues in a similar fashion that 'predicative propositions' and 'imperative

propositions' involve combining properties with objects in different ways, which allows for distinguishing between indirect reports like "Jones told Smith that Smith will go to the store" and "Jones told Smith to go to the store" (2007, 150).

For the purposes of this paper, I will henceforth simply assume that there are imperative contents and not just imperatival force, without committing to any of the above ways of cashing out this idea. This assumption opens up the possibility of an alternative intentionalist strategy: perhaps affective experiences feel the way they do in virtue of their imperatival content.

Why be an imperativist? Colin Klein (2015) begins with homeostatic sensations like thirst, hunger, and itching, which have as their biological function restoring some sort of bodily balance. They plausibly don't do this by informing us about the source of the imbalanced condition so as to inform potential deliberation. Rather, they function like a fire alarm does: they insistently tell us what to do – Drink! Eat! Scratch! – hogging our attention and interrupting other projects. Still, they're not mere reflexes: thirst isn't a stimulus to which you automatically respond to by drinking. In order to play just this kind of functional role, Klein argues, homeostatic sensations must have imperatival content and consequently a corresponding nagging phenomenal character. And he holds that much the same goes for pains, whose imperatival content gives them their *sensory* phenomenology.

Importantly for my purposes today, imperativism has also been proposed as an account of *valence*. The main argument for it appeals to the motivational role of valence:

1. The phenomenal character of an experience matches its intentional content.
2. Valenced phenomenal character is inherently motivating – when we take the experience at face value, we're motivated to act.
3. For taking an experience at face value to motivate us to act, the intentional content of the experience must be imperatival rather than indicative.

4. So, valenced phenomenal character consists (at least partly) in having imperatival content (stronger: affects have valence in virtue of their content containing a command).

Premises 1 and 2 are both widely accepted. For premise 3, imperativists rely on the sort of Humean considerations that I mentioned in criticizing evaluativism. For a content to be such that it directly motivates us when accepted or taken up, it must have satisfaction-conditions rather than truth-conditions. If we accept this, the ‘pushy’ phenomenology of valenced affects requires them to have imperatival content, possibly in addition to other kinds of content.

A number of different versions of imperativism have been proposed, in particular in the context of theories of pain. I’ll introduce them by considering what they say about a situation in which you’re in pain because you’ve twisted your ankle. According to *first-order imperativism about valence* defended by Manolo Martinez (2011), the pain sensation that you feel has two kinds of content. First, its *indicative* content says something like “There’s damage to your ankle”. Having this kind of content, according to him, accounts for the sensory phenomenal character of your ankle pain. Second, the sensation has *imperatival* content, which says something like “See to it that the damage in your ankle goes away!”, and accounts for the negative affective valence of pain. This is a first-order view, because the valence-generating imperative is directed toward a non-mental state of affairs.

First-order imperativism has been subjected to various criticisms. Evaluativists, naturally enough, complain that it can’t account for the reason-giving force of pain (Bain 2011), or variations in intensity (Cutter and Tye 2011). I’ll come back to these points in Section 3. But it’s worth noting that first-order views also face criticism within the imperativist camp. In particular, Luca Barlassina and Max Hayward level a number of charges against it. First, they hold that it is not sufficient for negative valence that experience

contains a first-order imperative (2019, 10). For example, thirst tells us to drink, but mild thirst isn't unpleasant. I don't think this is a serious problem, however – saying that mild thirst is mildly unpleasant is not much of a bullet to bite. Second, however, Barlassina and Hayward hold that there can be valence without a world-directed imperative. Depression is unpleasant, but, they claim, according to best theories, it's characteristic of depression that “world-directed urges have disappeared” (ibid.). And third, they claim that first-order internalist views haven't spelled out how to distinguish between imperatives that give rise to negative valence and those associated with positive valence. They point out that it won't do to say that negative imperative content “directs its addressee to stay clear from, or to avoid, or to get less of, something in the non-mental world” (11) – after all, agonizing hunger motivates us get more food. One could say that hunger tells us to stop having an empty stomach – but, Barlassina and Hayward note, we need a principled basis for distinguishing between sensations whose contents are characterized by “Stop having an empty stomach!” and “Eat something!”. Finally, they maintain that first-order imperativism can't explain why we should get rid of, say, unpleasant pain as well as damage, if the imperative concerns the latter. I'll address the last three problems at various points below, since my own view will be close to Martinez.

Higher-order imperativism, defended by Colin Klein, would avoid these problems. According to it, the unpleasantness of pain is instead “constituted by a second-order imperative directed toward the first-order sensation” (2015, 186). Here the content of the imperative can be glossed as “Don't have that sensation!”. This type of view makes sense of the fact that the unpleasantness of pain motivates us to take painkillers, even though we know perfectly well that they won't help with damage to our ankle, for example. But it also has serious costs – for example, it's implausible both that a neutral sensation becomes unpleasant in virtue of a command to get rid of it, or that the command to get rid of a sensation is itself

unpleasant (Barlassina and Hayward 2019, 16-19). To avoid these issues, Barlassina and Hayward advocate instead what they call *reflexive imperativism*, according to which, roughly, the imperatives involved in unpleasant experiences tell us to get rid of the experience they're a part of, while other experiences are pleasant in virtue of containing imperatives to pursue more of the very same experience. I'll come back to this view in Section 3, but to foreshadow a little, its key weakness is that it is forced to give a strained explanation of the world-directed intentionality of affective experiences, unlike first-order imperativism.

2. Generalizing Imperativism About Valence

In this section, I'll begin to develop a generalized version of imperativism about valence that is intended to apply to all kinds of affective state. Put in crudest terms, the view is that an experience has negative valence when it simultaneously presents things as being in a certain way and commands us to ensure that they won't be that kind of way.⁴ Here is a more precise statement:

Relational Imperativism

An experience has positive (negative) valence for S when it has a) indicative content *p* that represents an actual or possible state of affairs that bears on a (possibly novel) background concern of S and b) subjectively authoritative imperatival content telling S to see to it that states of affairs relevantly like *p* will (won't) obtain in the future by engaging in some behaviour characteristic of the affect type.

⁴ Relational imperativism is thus parallel to those externalist views that hold that unpleasant experiences consist in a relational state of taking things to be in a certain way and intrinsically wanting them not to be that way (Jacobson 2018) and, in a different way, relational expressivist views about normative belief (Toppinen 2013, Schoeder 2013).

I will argue that this kind of intentional content accounts for the inherently motivating phenomenal character of valenced affects, and respond to objections to it (mostly) in the next section.

But first, I'll spell out my thesis in more detail. I'll illustrate it with the help of four kinds of negative affect:

(e₁) Pain: My ankle hurts.

(e₂) Sensory displeasure: I have the unpleasant experience of having bitten a rotten apple.

(e₃) Attitudinal displeasure: I'm displeased that Sweden beat Finland.

(e₄) Fear: I'm afraid of a bear in front of me on the path.

(I'll come back to moods and positive affects later.) Let's begin with the indicative content.

It's become common to think that the felt aspect of fear is itself a *feeling towards* the bear, for example (Goldie 2000). The natural way to cash this out is to say that the affective experience itself has this state of affairs as its indicative content. It's part of the affective experience that things appear to you to be in a certain way – even if you don't believe they are that way (say, in the case of phantom pain in a limb you know you don't have). In the cases above, the appearances might be as follows:

(a₁) Pain: My ankle is damaged.

(a₂) Sensory displeasure: There's a rotten taste in my mouth.

(a₃) Attitudinal displeasure: Sweden beat Finland.

(a₄) Fear: There's a bear poised to do something I really wouldn't like to me.

Note that some of these appearances concern the world, and some concern mental contents. In particular, sensory displeasures represent sensations (naturally enough). In each case, you could have the feeling even if the appearance isn't veridical.

In the case of valenced affective experience, the state of affairs represented by the experience must *matter* to the subject. Bearing in mind that it is at least conceptually possible for many sorts of creatures to have valenced experiences, it is best to construe the relevant sense of mattering as thinly as possible. I'll thus say that the indicative content engages with some *background concern* of the subject. By background concern I mean an open-ended standing desire in the broadest sense. It could be the biological drive to survive or avoid bodily damage, or it could be a fan's dispositional desire for her team to win, or endorsement of a moral standard. In my examples, the relevant concerns might be expressed as follows:

(bc₁) Pain: May there be no bodily damage to me!

(bc₂) Sensory displeasure: May I not have the taste of spoiled food in my mouth!

(bc₃) Attitudinal displeasure: May the Finnish team do well!

(bc₄) Fear: May I not get hurt!

The role of such background concerns in affect is most obvious when they differ among people – for most people in the world, it's neither here nor there whether Sweden beats Finland or not, so they have no affective response to such news. Because of such facts, psychologists working on emotion tend to emphasize the importance of usually automatic and non-conscious *appraisal* of goal-relevance for the generation of affect (e.g. Frijda 1986). Some, like Robert Roberts (2003) even hold that negative valence consists in appraisal of something as frustrating one's concerns, but this goes too far. First, non-conscious appraisal is not sufficient to account for valenced *phenomenal* character. This is not to say we can't have, say, negative emotions without realizing it – but insofar as such emotions are occurrent, they

shape the horizon of our experience while escaping our attention, similar to the way stress functions (Haybron 2008). And it's not plausible that negative valence consists in consciously appraising something *as* concern-frustrating, as Deonna and Teroni (2020) observe. Second, *prior* concern or appraisal is not necessary for valence. Sometimes our (non-valenced) experiences explain why we have the concerns we do. For example, a first-time experience with spicy food may immediately yield a novel liking. I'll return to this soon when discussing subjective authority.

The second and distinctive kind of content for imperativism is the first-order, world-directed imperative. Recall here Martinez's characterization of the imperative in pain as "Don't have this bodily disturbance!". For the general case, it's too narrow to specify the target of the imperative as the indicated state of affairs itself – after all, it might well be an unrepeatable event that will in any case not occur again. It's thus better to cash out the imperative as "See to it that *the sort of thing* indicated in this experience won't obtain in the future!", where what counts as "the sort of thing" depends on the context. Note that contrary to Barlassina and Hayward's characterization of first-order views, the imperative doesn't just tell us to get less of *something*, but specifically refers to the indicative content of the experience. In their example of agonizing hunger, the indicative content is that we lack nourishment, and the imperative content tells us to see to it that we won't lack nourishment any more, even if we have to sacrifice a lot of other things. There's nothing arbitrary about this *relational* structure giving the experience negative valence.

Before giving examples, I want to emphasize, first, that the claim isn't that something verbalized along these lines is going through the subject's head. Rather, the formulation is meant to capture the satisfaction conditions of an experiential content that is itself non-conceptual in the sense that it doesn't require possessing a concept of the future, not to mention the concept of being indicated in experience, among others. Stipulating such contents

is needed to capture the dynamics of experience and action – in the very simplest case, feeling bad in a certain way tends to motivate a certain type of action, which tends to make the bad feeling go away if successful. They should also make sense of the bodily action readiness characteristically associated with many valenced experiences. Second, my verbal characterization is meant to give the *form* of the first-order imperative central to negatively valenced experience. On my view, many particular negative affects tell us to achieve this type of outcome in a particular way – after all, fear doesn't feel bad in the same way as anger, and motivates different behaviour when taken at face value. So I would suggest that pain tells you to see to it that the sort of bodily damage indicated by the experience won't obtain in the future *by protecting the damaged body part*, and fear tells you to see to it that the sort of state of affairs indicated in its content won't obtain *by seeking safety*. The specific actions they motivate may depend on the context and the agent's background beliefs.

An important aspect relational imperativism is the *authority* of the imperative for the subject. Pain or fear is more like a boss telling you what to do than like a passer-by making a request. What could be the source of such subjective and practical authority? Not, for example, that it would be unpleasant if we didn't do as commanded, since that assumes what we're trying to explain (Bain 2011). Fortunately, there is a candidate at hand: the background concern. When fear tells you to run away, it speaks *in the name of* your background concern to stay alive and unharmed. Less metaphorically, the suggestion is that our motivational and affective systems interact, so that only affects that link up in the right way with background motivation have valenced character (which is not to say affects that don't do so lack phenomenology altogether). As I noted above, this is compatible with experiences *generating* the motivation. It could be that the first time you taste spicy food, the key content of your experience can be verbalized as "There's a fiery sensation in my mouth; let me see to it that I have more sensations like this!". If this experience engages your motivational system, say by

generating a (possibly short-lived) liking for this sort of thing, your experience will have positive valence, even if you didn't have a *pre-existing* concern that was satisfied by spicy sensations.⁵

One reason to adopt this model is that it makes good sense of the curious phenomenon of *pain asymbolia*, where people with certain sort of brain damage report feeling pain when in response to certain stimuli, but lack any motivation to avoid it and deny that it's unpleasant. Nor do they act to avoid threats to their bodily integrity. Both Colin Klein (2015) and David Bain (2014) have proposed that pain asymbolia is best explained in terms of absent concern for the condition of one's body. Bain argues convincingly that what happens in such cases is that the lack of concern alters the phenomenal character of pain such that it is no longer unpleasant for the subject. He claims that imperativism can't explain "why not caring about one's bodily integrity should prevent pain's imperative content". Evaluativism, for its part, holds that if I lack concern for my body, I won't find the damage indicated by pain as bad for me, so it isn't unpleasant. However, if valenced character is linked with subjective authority of imperatives, which requires a background concern, imperativism also predicts that the asymbolics' pain is not unpleasant for them in virtue of the absent concern for their bodily integrity. It is, in fact, much more natural to link background concern with the authority of imperatives than lack of evaluation – if evaluation is straight representation, why would not caring about whether something happens entail that I don't think that it is *bad* for me? What would stop me from thinking "That's bad for me, but I don't care about it"? The imperativist account thus seems actually superior to the evaluativist one on this score.

With these considerations in mind, here's what the imperatival content of my examples might look like:

(i₁) Pain: See to it that your ankle isn't damaged by protecting it!

⁵ You might have a standing desire to eat food that tastes good, to be sure – but that couldn't explain why spicy food *tastes good* for you.

(i₂) Sensory displeasure: See to it that you will not have this sort of sensation in the future by avoiding what caused it!

(i₃) Attitudinal displeasure: See to it that Sweden won't beat Finland in the future!

(i₄) Fear: See to it that there won't be a bear poised to do to you something you really wouldn't like by seeking safety!

It is important here that it doesn't follow from an affect telling you to act in a certain kind of way that you will be motivated to act like that – your desires, plans, and beliefs might prevent the *conative uptake* of the imperative. Feelings don't necessarily motivate, even if the motivation is non-instrumental when it's there. And when there is conative uptake, the specific direction it takes depends on the subject's background beliefs – whether you rub the ankle that hurts or lift it in the air depends on what you think will help protect it. This is why it's a mistake to identify imperatival content with 'urges', if they're understood as motivational states. This is crucial for understanding attitudinal displeasure in imperativist terms, since it is often not possible to do anything about what displeases us (or what we regret or grieve). Knowing this, we may not be motivated to do anything either – though in some cases, we engage in a kind of magical thinking and are motivated to, say, yell at the TV as if it changed the way our team plays. This is all perfectly consistent with the displeasure telling us to change the way things are – indeed, the felt inability to respond to an experienced command does nothing to make us feel better.

So, the view holds that fear, for example, has a negatively valenced character in virtue of (a₄) *There's a bear poised to do something I really wouldn't like to me* and (i₄) *See to it that there won't be a bear poised to do something you really wouldn't like to you by seeking safety!*, when the latter has authority in virtue of (bc₄) *May I not get hurt!*, which (a₄) is appraised to (possibly) frustrate the way things are going. Before assessing the account, let's

take a look at some positive affects as well. They pose a prima facie challenge for imperativism, since it has often been observed that positive emotions are typically less directly linked to specific actions than negative ones. Still, positive emotions do tell us to do something. The best way to cash it out seems to me to be along the lines of “Explore further possibilities for realizing or relishing this sort of thing in the future!” This type of content fits well with the broaden-and-build evolutionary rationale suggested by Barbara Fredrickson (2001) for positive emotion. In the ordinary case, positive feelings result from its seeming to us that things are going our way, so it’s no surprise they don’t call for taking things in a different direction. Rather, they tell us to let things be (and indeed prevent them from changing) and look around for more of the same at one or another level of abstraction. I’ve also included relishing in the content of the imperative, since sometimes further world-directed action directed at realizing the state of affairs is superfluous or impossible, but we can still engage in *mental action* of seeking ways to appreciate the situation in further ways.

Here’s how some paradigm positive feelings look like from this perspective. Start with *sensory pleasure*. Suppose you bite into a ripe persimmon. The intentional content of your experience might be “There’s a sweet taste on your tongue; explore further possibilities for this kind of taste in your mouth and attend to its characteristics!” (More generally, sensory pleasures call for more of the same sort of sensation by way of same sort of stimulation.) A paradigm *attitudinal pleasure* might be taken in your favourite team winning. Its content would be along the lines of “Your team won; [explore further possibilities for this kind of thing, and] attend to this fact!” As in attitudinal displeasure, the action directive in such cases is typically idle, since you neither can nor need do anything about the fact that pleases you. But you can pause and contemplate what happened, and perhaps bring to the attention of others. *Joy* can be quite similar, though perhaps more intense, and call for more energetic

exploration of similar possibilities. And *interest* informs you that its object has so far aroused and satisfied a desire to know, and tells you to explore it further.

So far, I've set aside an important kind of valenced experience, moods like (occurrent) anxiety, irritability, depression, tranquillity and elation. (Note that while we sometimes use these terms for dispositional conditions, they clearly can also denote modes of consciousness.) They pose a *prima facie* challenge to first-order imperativism, since like positive emotions, they don't seem directly linked to action. Indeed, in some moods action doesn't seem possible, as in a kind of depression, and in others it doesn't seem necessary, as in tranquillity. Nevertheless, I think the present account has sufficient resources to explain what's going on. Say that I'm anxious about my coming move. The experience involves associating the move with some unspecified threats. It's negative in virtue of telling me to get rid of those unspecified threats. Since I don't know what to do about unspecified threats, I can't respond to such an imperative with specific motivation (if not with motivation to cancel the move altogether, which is not available if anxiety is more general). I can respond with bodily action readiness to fight or flee whatever comes my way, however, which is a typical sign of anxiety. In the case of some sorts of depression, to be sure, even such action readiness (as well as motivation) can be absent. Insofar as it is not simply a state of suppressed affect, it could be because depression represents us as pointlessly trying to do things that are not worth it, whatever we do, and tells us to see to it that we don't engage in further pointless activity. On the positive side, tranquillity has as its indicative content something like the thought that things are going all right without my interference. Applying the relational formula, the mood's imperative content tells me to see to it that this will be the case in the future and to engage in the mental action of appreciating the way they are right now. This, again, seems to have just the right motivational consequences: according to it, conative uptake of tranquillity

motivates not to mess up anything by interfering with it and to contemplate the niceness of our situation, so we hang back and smile.

Having filled in some details of relational imperativism, we're now in a position to tentatively evaluate it. Does imperativ content of the kind defended account for the phenomenal character of valenced affects? Some find it in general hard to believe that any imperatives would explain valence. Other think it is evident that they do – Martinez says that “This is what disgust feels like: it feels like being compelled not to get polluted by disease-causing substances” (2015b). Here introspection may yield different answers. This is a difficult question to adjudicate, since the rules of the game are not at all clear. That's why I have focused instead on the functional role of valenced character. And here generalized imperativism does seem to provide a plausible, nuanced picture of how both positive and negative affects are inherently if contingently motivating. If, say, interest tells you to explore its object (or the sort of thing its object is) further, it's easy to see why yielding to your interest in, say, a Lego set gets you to play with it and be on the lookout for more Lego sets, and consequently endlessly talk to your parents about getting new Lego sets. Sensory displeasure all on its own tends to motivate you to avoid the kind of food – or sound or texture – that caused it, if you conatively assent to what it tells you. And pain, among other things, gets us to avoid stepping on a hurt ankle. So generalized imperativism seems to do a good job when it comes to how valence motivates us to engage with the world. But there are, of course, serious challenges to it, as we'll see next.

3. Objections and Replies

Objection 1: Imperativism Can't Explain How Affects Rationalize

One might grant that relational imperativism does a good job of explaining the motivational role of valence, but what about *rationalizing* action and judgment? One of the main

arguments for evaluativism was that it can explain this – as Bain puts it, “unpleasant pains rationalise avoidance behaviour in virtue of presenting certain bodily states as bad” (2013).

Carruthers suggests that felt badness is an element of the phenomenology of emotions in particular:

When a bear looms out of the bushes while one is hiking (causing fear) it is the threatening aspect of the bear (its size, its claws) that seems bad. All of one’s focus when afraid is generally outward-directed, targeted on the object of one’s fear. [...]

Negative valence represents the presence of the bear (or its threatening aspect) as bad. (2018, 664–665)

Is there any way for imperativists to make sense of the rationalizing role of affects and their presentational phenomenology? I think so. The first step of the argument is reformulating the target in a way that doesn’t beg any questions in favour of evaluativism. Consider, first, how Christine Korsgaard characterizes animal experience:

We might suppose that when an antelope flees from a lion, she is afraid. But she need not think about the fact that she is afraid ... Perhaps she just sees the lion as a danger, as a thing that is *to-be-fled*. (Korsgaard 2018, my emphasis)

For Korsgaard, affective experiences involve what she calls *teleological perception* of our environment. Such perceptions suffice to rationalize flight just as well – or actually better – than a generic perception of badness (after all, it’s a deliberation-inviting open question what to do with something bad). In emphasizing teleology in connection with emotions, Korsgaard joins up with a long tradition. Consider Sartre’s description of pity or sympathy:

I feel pity for Peter and I come to his aid. For my consciousness, one thing alone exists at that moment: Peter-having-to-be-aided. This quality of ‘having-to-be-aided’ is to be found in Peter. (Sartre 1936/2004, 10)

I’m going to label the sort of appearance that Korsgaard and Sartre talk about an *affective practical seeming*. It is a species of non-doxastic appearance: it can affectively seem to us that things are in a certain way even if we don’t have a corresponding belief (and even if we have an opposing one). This kind of seeming is what imperativists must explain to explain how valence rationalizes. Why is it that when we’re afraid of something, it appears to us as something to be fled? To be precise, the datum isn’t that affective practical seemings are somehow inherent in valence; rather, it suffices if imperativists can explain why experiences with imperatival content predictably give rise to affective practical seemings. This is important, because such seemings clearly couldn’t *consist* in imperatives – they have a mind-to-world direction of fit.

What I want to argue is that relational imperativism offers a natural explanation of affective practical seemings, so that we don’t need to be evaluativists to account for them. The story is simple and appeals to a widely recognized psychological mechanism of *projection*. The *locus classicus* for this kind of line is Hume’s famous observation: “taste has a productive faculty, and gilding and staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation.” J.L. Mackie (1977) develops the view further, arguing that we have a tendency to see commands issued by socially accepted norms or purported divine authority as inhering in actions themselves, and Simon Blackburn (1993) explains why projecting our attitudes in this manner serves useful purposes. For example, instead of focusing on what someone told me to do, I can focus on whether the thing itself really is to be done.

I argued above that when we have a valenced affective experience, it is as if we're being authoritatively told or invited to do something. If there is a psychological mechanism of projection, as there seems to be, it is predictable that such demands change the way our situation appears to us. For a parallel, consider the situation when an officer tells a private to load his gun. Directly, the order just concerns the action. But we can also ask how horizon of practical possibilities with respect to the gun looks like to the private who accepts this order. The gun will appear to him as something *to-be-loaded* when he thinks about what to do. Similarly, then, if my fear has an indicative content (e.g. "that bear might attack me") that bears on my concern (e.g. for personal safety), and related authoritative imperative content (e.g. "see to it that sort of situation in which the bear might attack me won't obtain by fleeing it"), it will appear to me teleologically as something that summons me to leave the scene – as something *to-be-fled-from*, at a certain cost to other things I care about.

Suppose that imperativism can explain why valenced experiences (at least tend to) include affective seemings. The rest of the rationalization story is quickly told. As I look at the bear, it appears to me as something *to-be-fled*, and so as something I have reason to flee. From my perspective, this suffices to rationalize flight, or at least give rational support to the action, even if it is outweighed by other considerations. It is perhaps more interesting that this story also makes possible a kind of *reconciliation with evaluativist views*. After all, recall my worry about what it is to non-conceptually represent something as bad or good. Now we have a good candidate answer: perhaps it just consists in representing it as summoning less or more of the same sort of thing – something non-conceptually appears as bad to a certain degree in the light of an objectified imperative, if you will, saying that it is to be avoided at such-and-such cost. In that case, my view is compatible with some evaluativist accounts – but gives a

further analysis of their central content.⁶ This seems to me like a good further argument in favour of relational imperativism over evaluativism.

Objection 2: Imperativism Can't Make Sense of the Intensity of Valence

One standard objection to imperativism is that it can't capture the *intensity* of affective valence, the fact that some pains or fears feel worse than others (Cutter and Tye 2011). The argument is simple:

1. Some pains, fears, and hopes are more intense than others – intensity is a core feature of valenced phenomenal character.
2. Commands can't be more or less intense (even if they can be, e.g. more or less polite).
3. So, there is no way to explain intensity in terms of commands.
4. So, imperativism can't account of valenced phenomenal character.

Klein and Martinez have offered a model for rejecting premise 2. Their solution begins by stipulating that the content of an imperative isn't just its satisfaction conditions, but also includes a ranking of all possible worlds. They propose that like the Cold War era Autovon telephone system, whose calls included a signal indicating their relative priority (so that, for example, Flash calls rank their connection as more important than Routine calls, and Routine calls rank their connection as less important than Flash calls), pain commands, too, rank their satisfaction relative to each other. In the simplest case, imperative *i* is more forceful than imperative *j* when not only the ranking function of *i* ranks the worlds in which *i* is satisfied over those in which *j* is satisfied, but so does *j*'s ranking function – *j*, as it were, yields to *i*. (For a more sophisticated model, see Martinez 2015a.) Correspondingly, a more intense affect issues a more forceful command in this sense: “So, for example, one pain is more intense than

⁶ This claim is obviously parallel to the claims that expressivists like Blackburn 1998 and Gibbard 2003 make regarding non-naturalism. There is a lot more to say on this point, but space will not allow me to do it here.

another just in case their world-rankings agree that one should tend to the body part involved in the first pain at the expense of the one in the second.”

My worry about this attempt is that this only works for very special cases, in which imperatives make mutual reference to one another (or third ones, as in their more sophisticated model). It is highly implausible that this would be a necessary condition for experienced imperatives, which very simple creatures can experience, after all. Fortunately, there seems to be a simpler imperativist account of intensity available. In their informal description of intensity, Klein and Martinez refer to a difference in opportunity cost that one is willing to bear for more or less intense pain. What I want to propose is that that we can make use of the same notion in defining the force of a command. In simplest terms, one command is more forceful than another when it requires us to bear a greater opportunity cost to carry it out. Compare these three military orders:

(1) Hold the line!

(2) Hold the line at any cost!

(3) Hold the line at the cost of some but not nearly all of our soldiers dying!

All of these commands have the same satisfaction conditions: they are satisfied when the soldiers hold the line. But (2) is more forceful than (1), since (1) implicitly allows some leeway – it leaves open that there’s some costs too large to bear to hold the line. Taken literally, (2) ranks all the worlds in which the imperative is satisfied higher than any of the worlds in which it is not satisfied, regardless of consequences for anything else. In reality, even orders like (2) express commands that are not absolute, so that there are some costs whose avoidance is in the spirit of the command even if it’s not in the letter, such as destruction of all humanity. In rational armies, it’s mutually understood that (1) expresses

something like (3). Exactly which ranking function it expresses – which losses go too far – is a matter of context.

The difference between (2) and (3) comes out in the way they rank the following possibilities (for example):

(w_1) The soldiers hold the line & some soldiers die or are wounded (up to -5 utils)

(w_2) The soldiers hold the line & all soldiers die in the course of doing so (-10 utils)

(w_3) The soldiers don't hold the line & all soldiers escape damage (0 utils)

Assuming that each command tacitly also contains “Sacrifice as little as possible!”, (2) ranks the possibilities as $w_1 > w_2 > w_3$. (3), in turn, imposes a ranking of $w_1 > w_3 > w_2$. In other words, (2) tells the addressee to hold the line even at the cost of -10 utils rather than fail to do so, in spite of the upside of there being no damage. (3) tells the address to hold the line at the cost of up to -5 utils rather than escape with no damage (or, for that matter, escape with minor damage), but it also ranks escaping to live to fight another day higher than everyone dying. Thus, (2) is more forceful than (3), because it demands to bear a higher opportunity cost for complying than (3) does. Here is a general formulation of the thesis:

Forcefulness of Imperatives

Imperative i is more forceful than imperative j iff i ranks worlds in which i is satisfied and the addressee loses up to u utils higher than worlds in which i is not satisfied and the addressee doesn't lose up to u utils, while j ranks worlds in which j is satisfied and the addressee loses up to v utils higher than worlds in j is not satisfied and the addressee doesn't lose up to v utils, and $u > v$.

Again, the contrast with Klein and Martinez's view is that my model doesn't assume that the ranking conditions of imperatives make reference to other imperatives, but just that imperatives at least implicitly tell us to bear some level of opportunity cost.

With the account of comparative forcefulness of imperatives in hand, accounting for the intensity of affect is straightforward: the more forceful the imperative, the more intense the affect. For example, the more forceful the order to explore more possibilities for the same sort of thing, the more positive the valence of the affect is. So, a more intense joy, for example, tells you to sacrifice more of other good things to explore more possibilities for the sort of thing indicated in its content than a less intense one.

Objection 3: Shooting the Officer

My version of imperativism, like first-order imperativism in general, holds that affects have valence in virtue of telling us to do something about our situation. But, a well-known objection goes, how could first-order imperatives explain how they also motivate and rationalize action that aims to do something about our experience itself? For example, why take painkillers? We do, after all, have reason to get rid of even non-veridical pains, like pain felt for a phantom limb, or pain caused by something that doesn't in any way damage us (Bain 2011).

This objection is evidently parallel to what is known as Shooting the Messenger Objection to evaluativism. As Hilla Jacobson puts that objection in the context of pain, "If we endorse the suggestion that the first-order con-attitude which is constitutive of pain is cognitive, we are committed to a picture according to which each time we take a pain killer, we behave like the ruler who killed the messenger who brought him bad news." (Jacobson 2013, 519) If unpleasant pain just tells us that something is bad for us, it isn't clear why it would itself be bad, or why we'd have reason to get rid of it. Since imperativism holds that

pain is more like an officer than like a messenger, I'll label the corresponding challenge the Shooting the Officer objection:

Shooting the Officer

1. According to relational imperativism, negative valence is constituted by a command to see to it that the indicated content won't obtain.
2. There is nothing non-instrumentally bad about being commanded to see to it that the sort of thing indicated in one's experience won't obtain, and no reason to get out of the state of being so commanded.
3. So, relational imperativism can't explain why we have reason to get rid of negative affective states by e.g. taking painkillers.

My response to this objection takes the same form as Bain's response to Shooting the Messenger. The core problem with the argument is that the second premise begs the question against imperativism (as well as evaluativism). After all, suppose that imperativism is right. In that case, pain's unpleasantness just is a matter of having a certain sort of relational imperative content. Therefore, if unpleasantness is intrinsically bad for you, being commanded to see to it that the sort of thing indicated in one's experience won't obtain is intrinsically bad for you. Bain (2019) calls his corresponding case for evaluativism the "Leibniz's Law Argument". I see no reason why an imperativist couldn't use exactly the same line. Thus, the reason why unpleasant pain is bad, and why there's a reason to get rid of it, is that it is unpleasant.

There is, to be sure, a further problem for relational imperativism. Even if it is as a matter of fact non-instrumentally bad for you to undergo an unpleasant experience, that doesn't suffice to rationalize wanting to rid of it, unless you also conceive of it *as* bad. Indeed, so far we still lack an explanation for why we even *want* to get rid of unpleasantness,

as opposed to damage or danger or loss. This is indeed a challenge for first-order views. Barlassina and Hayward formulate the putative explanandum instructively by claiming that affective experiences have a *reflexive* motivational force: “The mental state which affective phenomenal character motivates us for or against is the very same mental state that has affective phenomenal character.” (2019, 5) This is a key motivation for their own reflexive imperativism, according to which unpleasant sensations basically tell us to get rid of *themselves*, or more precisely the experience they’re a part of, saying “Less of me!”.

I think the right response begins by denying reflexive motivation. It’s not the case, for example, that an unpleasant experience U motivates us to avoid or get rid of U itself, or even a state like U. Take a world-directed emotion like grief. It is surely a negative affective state. But it doesn’t intrinsically motivate us to take grief-killers. What we wish for in the first instance is for the world to be different – for the loved one we lost to come back to us, and not just as a means to make ourselves feel good. Normally, of course, we don’t think we can do anything, but people who believe in communion with the dead are sometimes motivated by grief to take part in seances. For a more mundane case, fear will again suffice – it tells us to flee, not to take a pill that gets rid of itself, even if one becomes available.⁷ It may be less obvious that not even sensory pleasures or displeasures motivate us to pursue or avoid pleasures or displeasures. But that’s the case nevertheless: the pleasure of eating a perfect strawberry doesn’t motivate me to pursue more *pleasure*, but to *taste more strawberries*, that is, to seek more of the pleasing sensation.

This is not to say that we don’t want to avoid unpleasant experiences. We surely do – but that’s different from saying that unpleasantness itself tells us to get rid of itself. After all, we desire to avoid unpleasantness even when our experience is currently pleasant. It’s rather a

⁷ For Barlassina and Hayward, valenced affects motivate us to pursue worldly aims as a *means* of improving our hedonic state, which serves as reward or punishment (Barlassina and Hayward 2019, 28). If that were the case, we should be motivated to take a pill that gets rid of fear if we believed it to be more effective means than fleeing the danger.

background concern that we (contingently) have (cf. Cutter and Tye 2014, who also offer a plausible evolutionary explanation). And it helps explain why we sometimes do have unpleasantness-directed motivation. Suppose I'm undergoing an unpleasant experience, such as a toothache. The toothache itself tells me to get my tooth fixed (and indeed motivates me to find a dentist). But I'm in addition also *attitudinally displeased* by the fact that I'm undergoing this unpleasant experience. As above, this attitudinal displeasure involves an indicative content, *I'm undergoing an unpleasant experience*, which I appraise as frustrating my background concern *Let me not undergo unpleasant experiences!*. This background concern lends authority to the imperatival content *See to it that you won't be undergoing an unpleasant experience like this in the future!* (Note that unlike higher-order imperativism, the present account can avail itself of this direction of explanation, because we're not now trying to explain unpleasantness, but motivation to get rid of it.) Conatively taking up *this* second imperative, then, explains my motivation to take painkillers. And because the unpleasantness of the experience will also appear to me as to-be-avoided as a result of projecting the authoritative imperative, it also subjectively rationalizes taking painkillers.

On this story, then, it's being displeased with the unpleasantness of pain that explains and rationalizes taking painkillers, not the unpleasantness itself. It's worth noting that though pretty much everyone has the background concern to avoid unpleasantness, we're not always displeased to undergo unpleasant experiences. This is the case with grief and fear (at least as long as it's felt to be functional). Even pain can belong in this category, say for athletes who want to push themselves (Bradford ms). Unlike reflexive imperativism, relational imperativism then predicts, correctly, that we lack inherent motivation to get rid of the experience (as opposed to changing the world).

Conclusion

It is not an unappealing idea that what it is to be conscious of something as valuable in a particular way is to experience it as inviting or calling for pursuit or engagement or exploration or consumption, depending on its nature. I've argued that the motivational and rational role of such experiences gets neatly explained, if we take them to involve both a representation of how things are now and a related subjectively authoritative command, which also account for the teleological appearances. Valenced affective experiences are in this sense primitive modes of consciousness of apparent value, and quite possibly crucial building blocks for conceptually articulated evaluative thought.⁸

References

- Anscombe, G. E. M. (1957). *Intention*. Harvard University Press.
- Bain, David (2011). The Imperative View of Pain. *Journal of Consciousness Studies* 18 (9-10):164-85.
- Bain, David (2013). What makes pains unpleasant? *Philosophical Studies* 166 (1):69-89.
- Bain, David (2014). Pains that Don't Hurt. *Australasian Journal of Philosophy* 92 (2):305-320.
- Bain, David (2019). Why Take Painkillers? *Noûs* 2017 (2):462-490.
- Barlassina, Luca (ms). Beyond Good and Bad: Reflexive imperativism, not evaluativism, explains valence.
- Barlassina, Luca & Hayward, Max Khan (2019). More of me! Less of me!: Reflexive Imperativism about Affective Phenomenal Character. *Mind* 128 (512):1013-1044.
- Blackburn, Simon (1993). *Essays in Quasi-Realism*. Oxford University Press.
- Carruthers, Peter (2018). Valence and Value. *Philosophy and Phenomenological Research* 97 (3):658-680.
- Castañeda, Hector-Neri (1975). *Thinking and Doing: The Philosophical Foundations of Institutions*. D. Reidel Pub. Co..
- Chrisman, Matthew (2016). *The Meaning of 'Ought': Beyond Descriptivism and Expressivism in Metaethics*. Oxford University Press USA.
- Cutter, Brian & Tye, Michael (2011). Tracking representationalism and the painfulness of pain. *Philosophical Issues* 21 (1):90-109.
- Cutter, Brian & Tye, Michael (2014). Pains and reasons: Why it is rational to kill the Messenger. *Philosophical Quarterly* 64 (256):423-433.
- Deonna, Julien & Teroni, Fabrice (2015). Emotions as Attitudes. *dialectica* 69 (3): 293-311.

⁸ The stimulus for writing this paper came from some remarks made by David Bain in a conversation regarding my earlier and now superseded work on related issues (Kauppinen 2020). I'm also grateful to Uriah Kriegel, Lilian O'Brien, and Fabrice Teroni for written comments on earlier drafts, and Andrew Lee for some sharp challenges. Feedback from audiences in Houston, Clermont-Ferrand, and Helsinki was also very helpful.

- Deonna, Julien & Teroni, Fabrice (2020). Emotional Experience: Affective Consciousness and Its Role in Emotion Theory. In U. Kriegel (ed.) *The Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Evans, Gareth (1982). *The Varieties of Reference*. Oxford University Press.
- Frijda, Nico H. (1986). *The Emotions*. Cambridge University Press.
- Gibbard, Allan (2003). *Thinking How to Live*. Harvard University Press.
- Goldie, Peter (2000). *The Emotions: A Philosophical Exploration*. Oxford University Press.
- Hanks, Peter W. (2007). The Content–Force Distinction. *Philosophical Studies* 134 (2):141-164.
- Haybron, Daniel M. (2008). *The Pursuit of Unhappiness: The Elusive Psychology of Well-Being*. Oxford University Press.
- Heathwood, Chris (2007). The reduction of sensory pleasure to desire. *Philosophical Studies* 133 (1):23-44.
- Helm, Bennett W. (2002). Felt evaluations: A theory of pleasure and pain. *American Philosophical Quarterly* 39 (1):13-30.
- Jacobson, Hilla (2013). Killing the Messenger: Representationalism and the Painfulness of Pain. *Philosophical Quarterly* 63 (252):509-519.
- Jacobson, Hilla (2019). Not Only a Messenger: Towards an Attitudinal-Representational Theory of Pain. *Philosophy and Phenomenological Research* 99 (2):382-408.
- Kauppinen, Antti (2020). The World According to Suffering. In D. Bain, M. Brady, and J. Corns (eds.) *The Philosophy of Suffering*. Routledge.
- Klein, Colin (2007). An Imperative Theory of Pain. *Journal of Philosophy* 104 (10):517-532.
- Klein, Colin & Martínez, Manolo (forthcoming). Imperativism and Pain Intensity. In David Bain, Michael Brady & Jennifer Corns (eds.), *The Nature of Pain*.
- Klein, Colin, What the body commands : the imperative theory of pain. MIT Press.
- Korsgaard, Christine M. (2018). *Fellow Creatures: Our Obligations to the Other Animals*. Oxford University Press.
- Mackie, John L. (1977). *Ethics, Inventing Right and Wrong*. Penguin Books.
- Martínez, Manolo (2011). Imperative content and the painfulness of pain. *Phenomenology and the Cognitive Sciences* 10 (1):67-90.
- Martínez, Manolo (2015a). Pains as reasons. *Philosophical Studies* 172 (9):2261-2274.
- Martínez, Manolo (2015b). Disgusting smells. *Journal of Consciousness Studies* 22 (5-6): 191-200.
- Mendelovici, Angela (2018). *The Phenomenal Basis of Intentionality*. New York, USA: Oxford University Press.
- Millikan, Ruth G. (1984). *Language, Thought, and Other Biological Categories*. MIT Press.
- Mitchell, Jonathan (2018). On the non-conceptual content of affective-evaluative experience. *Synthese*:1-25.
- Pautz, Adam (2010). Do theories of consciousness rest on a mistake? *Philosophical Issues* 20 (1):333-367.
- Ridge, Michael (2014). *Impassioned Belief*. Oxford University Press.
- Roberts, Robert C. (2003). *Emotions: An Essay in Aid of Moral Psychology*. Cambridge University Press.
- Sartre, Jean-Paul (1936/2004). *The Transcendence of the Ego: A Sketch for a Phenomenological Description*. Routledge.
- Schroeder, Mark (2013). Tempered expressivism. *Oxford Studies in Metaethics*.
- Smuts, Aaron (2011). The feels good theory of pleasure. *Philosophical Studies* 155 (2):241-265.
- Tappolet, Christine (2016). *Emotions, Value, and Agency*. Oxford University Press.

- Teroni, Fabrice (2020). Valence, Bodily (Dis)Pleasures and Emotions. In David Bain, Michael Brady & Jennifer Corns (eds.), *Philosophy of Suffering*. New York: Routledge, 103-122.
- Toppinen, T. (2013). Believing in Expressivism. *Oxford Studies in Metaethics* 8.