

Forthcoming in *Ergo*

François Kammerer

Université catholique de Louvain

The Normative Challenge for Illusionist Views of Consciousness

Abstract: Illusionists about phenomenal consciousness claim that phenomenal consciousness does not exist but merely seems to exist. At the same time, it is quite intuitive for there to be some kind of link between phenomenality and value. For example, some situations seem good or bad in virtue of the conscious experiences they feature. Illusionist views of phenomenal consciousness then face what I call the *normative challenge*. They have to say where they stand regarding the idea that there is a link between phenomenality and value. If they accept that there is such a link, they might be committed to revisionary normative consequences (and some of them may prove to be uncomfortable). If they deny that there is such link, they might avoid revisionary normative consequences (without being guaranteed against them) but then they have to give reasons to deny that such link obtains, which is not a trivial task. The existence of the normative challenge does not show that illusionism is false, but it shows that illusionism might have important consequences in the normative domain, which have to be clarified.

Introduction

Here is an attractive thesis: there is a link between phenomenal consciousness and value. For example, some situations seem *good* or *bad* in virtue of the particular phenomenal states they feature. Think about someone being in pain: virtually everyone thinks that this situation is bad. Moreover, it is not only that we think that pain is bad; it seems bad *in virtue of its peculiar phenomenal feel*. We think pain is bad *because it feels awful*. Some philosophers have even

gone further in asserting the link between phenomenal consciousness and value. They have claimed that *nothing*, without phenomenal states, can have value. For them, phenomenality is a crucial ingredient to make something – anything! – good or bad.

Illusionists about consciousness claim that phenomenal consciousness does not exist, but merely seems to exist. That puts them in a difficult situation when considering the attractive view that there is a link between consciousness and value: they face what I call *the normative challenge for illusionist views of consciousness*. They have to say where they stand regarding the link between phenomenality and value. If they accept that there is such a link, they are led to revisionary consequences regarding what has value (more or less radical, depending on the kind of link they recognize). They then need to be clear about how far these consequences extend, and whether they can be limited. If, on the other hand, they deny that there is such a link, they might avoid revisionary consequences regarding what has value. However, they must then give reasons to reject this link, given that it is an attractive thesis. Moreover, even then, their views are not entirely guaranteed against *some* potential revisionary consequences regarding what has value, and these potential consequences ultimately need to be clarified.

Because each branch of the alternative raises difficulties, I take it that illusionists face a genuine challenge. However, that illusionism faces the normative challenge does not imply that it is *false*. Whether or not the normative challenge could ever constitute the basis of a robust argument against illusionism is in fact quite debatable. However, we must face up to the fact that illusionism encounters difficulties when one examines its consequences in the normative domain. Illusionists have to live up to this challenge by saying whether (and how) radical revisionary consequences can be avoided. If they cannot be avoided, illusionists have to explain how far exactly these consequences extend – and, ideally, how we can live with them. Besides, some philosophers seem to think that illusionism necessarily has disastrous revisionary consequences in the normative domain. These consequences appear to constitute one of their

motivations (though arguably not the only one) to reject illusionism.¹ This gives illusionists another reason to seriously discuss the consequences of their position in the normative domain.

I intend to show that illusionism *does not necessarily* have revisionary consequences in the normative domain. However, we would be wrong to think that illusionism can easily avoid *all* revisionary normative consequences. The most plausible ways to deal with the normative challenge probably lead to *some* revisionary consequences.

I first give definitions in order to explain what I mean by “phenomenal consciousness”, “illusionism” and “value” (§1). I then present two theses, designed to express two versions of the idea that phenomenality is linked to value (§2). I show why these theses are attractive (§3). I then present the normative challenge for illusionist views of consciousness (§4). I explore the main options open to the illusionist and I show that all of them create difficulties (§5-6). I conclude with some remarks about the consequences of the normative challenge (§7).

1. Definitions

Phenomenally conscious states (or “phenomenal states”, “conscious experiences”) are putative mental states endowed with phenomenal properties (or “*qualia*”), in virtue of which there is “something it is like” to be in them. Seeing a green patch, smelling a rose, feeling pain, are supposed to be typical examples of phenomenally conscious states. A subject who is able to enter into such states possesses phenomenal consciousness. Phenomenally conscious states must be notably distinguished, at least conceptually, from *access-conscious* states, defined as mental states bearing content available for use in reasoning, and rational control of speech and

¹ See for example what Galen Strawson writes about Dennett’s illusionism: “If [Dennett is] right, no one has ever really suffered, in spite of agonizing diseases, mental illness, murder, rape, famine, slavery, bereavement, torture, and genocide. And no one has ever caused anyone else pain [...] We must hope that [this idea] doesn’t spread outside the academy, or convince some future information technologist or roboticist who has great power over our lives” (Strawson, 2018).

action (Block, 1995). *Valenced* phenomenal states are “affective” phenomenal states, which are either *attractive* (positive valence) or *repulsive* (negative valence). Pain and sorrow are supposed to be typical examples of negatively valenced phenomenal states; pleasure and joy are supposed to be typical examples of positively valenced phenomenal states.

Illusionism about phenomenal consciousness (sometimes also called “eliminativism”² – here I do not distinguish between the two) is the thesis that phenomenal consciousness does not exist, but merely seems to exist. For illusionists (as opposed to *realists*) about phenomenal consciousness, we never enter phenomenal states and there is nothing it is like to be anyone. We simply enter states (introspective states) that incorrectly represent that we are in phenomenal states, thus creating the illusion of phenomenal consciousness. First versions of such views have been suggested in the 1960’s (Feyerabend, 1963; Rorty, 1965); this position has found prominent defenders in the late 1980’s (Dennett, 1988, 1991; Rey, 1995). It has been recently receiving more attention (Dennett, 2017; Frankish, 2016; Graziano, 2013; Pereboom, 2011).

Interest for illusionism has been growing fast in the last ten years, probably because it provides a robust (though counter-intuitive) defense of physicalism regarding the human mind – an otherwise very attractive metaphysical position, which has been threatened by arguments and thought experiments focusing on the alleged non-physical nature of phenomenal consciousness (Chalmers, 1996; Jackson, 1982). For many years, most philosophers of the mind have endorsed physicalism but have favored a *realist view* of consciousness. The most popular position has probably been the one David Chalmers dubbed “type-B” physicalism (Chalmers, 2002). According to this view, phenomenal consciousness reduces to physical processes, even though such reduction will forever remain somewhat opaque – in some views, to the point that we should expect consciousness to *seem persistently distinct* from physical processes (Aydede

² In the framework suggested by Irvine and Sprevak (Irvine & Sprevak, forthcoming), illusionism as I understand it would be equivalent to *entity* eliminativism (as opposed to *discourse* eliminativism).

& Güzeldere, 2005; Balog, 2012; Kriegel, 2009; Loar, 1997; Papineau, 2002; Sturgeon, 1994). This was supposed to give a way to maintain physicalism about consciousness, while accounting for the persisting difficulties encountered by all attempted reductions of consciousness to some physical process. However, there has been a growing concern that such views cannot ultimately succeed, and that it is impossible to conjoin physicalism about consciousness with a satisfying account of why consciousness seems persistently distinct from the physical (Chalmers, 2007; Demircioğlu, 2013; Goff, 2011; Levine, 2001, 2007; Nida-Rümelin, 2007). This point is still widely debated (Diaz-León, 2014; Elpidorou, 2013, 2016; Schroer, 2010; Shea, 2014), but the idea that we really can “have it both ways” (Loar, 1997, p. 598) – account at the same time for the ontological reducibility of consciousness and for its persistent epistemic irreducibility – seems somewhat less plausible than it once did. For that reason, illusionism, which allows us to defend physicalism about the human mind in a simple (but radical) way, by simply denying that the most problematic aspect of the mind (consciousness) really exists, has been receiving a renewed interest in the last years. Though it remains a minority view, it has found vocal support, both from proponents (illusionism has been described as the “obvious default theory of consciousness” (Dennett, 2016)) and from opponents (“if I were a materialist, I would be an illusionist” (Chalmers, 2018)).

A few things to note. First, what I describe here is what Keith Frankish (Frankish, 2016, p. 15-16) calls “strong illusionism”. Some philosophers have defended a position he dubs “weak illusionism”, according to which phenomenal consciousness exists, but does not have many of the properties it is usually thought to have (it merely seems to have these properties). My focus here is on strong illusionism (I will briefly go back to weak illusionism at the end of this paper). Second, illusionists think that there are no phenomenal states in reality. To talk about the *real* states and properties which are usually thought to be phenomenal, they talk of “quasi-phenomenal” states/properties (as suggested by Frankish, 2016, p. 15). A quasi-phenomenal

property is a non-phenomenal, physical property (possibly a wildly disjunctive property), plausibly a brain property, that we typically misrepresents as phenomenal – for example, through introspection. States that have quasi-phenomenal properties are quasi-phenomenal states. According to an illusionist, when I look at a red apple, I have a quasi-phenomenal perception of red, endowed with quasi-phenomenal redness; when I jam my finger in a door I enter quasi-phenomenal pain, etc. Third: in order to talk about what happens to me when I look at a red apple or when I jam my finger in a door in a way which remains *neutral* between illusionism and realism, I will talk of “allegedly phenomenal states/properties”. Allegedly phenomenal redness and allegedly phenomenal pain are properties and states that are phenomenal if the realist is right, but merely quasi-phenomenal if the illusionist is right. Fourth: in this paper, I will presuppose that, if we take the set of physical truths about the world as fixed, the existence *and* the non-existence of phenomenal consciousness are both conceivable. That is, that both zombies and genuinely conscious beings are conceivable, given physical truths about the world.³ So, if we are realists about consciousness, we think that we are genuinely conscious beings, but zombies are nevertheless conceivable. If we are illusionists, we think that we are zombies, but genuinely conscious beings are nevertheless conceivable.

Let us turn to value. I will presume that the bearers of value are *situations*: situations are good or bad. I will also talk of states (i.e. property-instantiations) which can compose situations as *grounding* value, when situations including these states have such value *in virtue of* these states being the case. A property, the instantiation of which grounds value, is a “value-making property”; the corresponding state a “value-making state” (A. Lee, 2018). A distinction is often made between *full* and *partial* grounding (Fine, 2012): A grounding B can mean either that B obtains entirely in virtue of A (full grounding) or that B obtains in part in virtue of A and in

³ At the end of this paper, I will say a word about views that reject this presupposition.

part in virtue of something else (partial grounding). When left unspecified, “grounding” here means full grounding.

We also have to distinguish between *intrinsic* and *extrinsic* value. Situations have intrinsic value when they have value in themselves, for their own sake, in a non-derivative way (Zimmerman, 2015). Pleasure, happiness, friendship, beauty and knowledge, are often cited as states that may ground intrinsic value (for an exhaustive list, see Frankena, 1973). Situations have extrinsic value when their value is derivative; for example, when they have value merely in virtue of the fact that they tend to bring about situations that have intrinsic value. When I talk about value without qualification, I mean *intrinsic value*. Philosophers sometimes also distinguish between various heterogeneous *kinds* of intrinsic value: for example, between epistemic value, ethical value and aesthetic value (Kriegel, forthcoming). I will not presuppose any such distinction here and will simply talk of *intrinsic value* (as something that may or may not be correctly thus specified), although my attention will be mostly focused on what people usually call “ethical value”.

The way I think about it, value can be *positive* or *negative*: situations are good when they have positive value, bad when they have negative value. I think values are *comparable* (some situations are better or worse than others) but I will stay neutral on whether or not they are *commensurable* (so that one situation can be, say, two times better than another one), though I suspect they can. I will also remain neutral on the relation between value and other normative notions such as duties and rights (when I will not, I will be explicit about it).

Finally, judgments about values are *normative judgments*. In this paper, I will make the (substantive) assumption that normative judgments *have a truth-value* (i.e. I will assume cognitivism regarding judgments of value). When adopting a new theory implies that we should revise our judgments regarding what has value (and/or the ordering of values), I will say that this theory has *revisionary normative consequences*. I will also assume – and this assumption

will not be used as a premise in any argument, but is important in order to grasp the general spirit of the paper – that we are *prima facie* resistant to revisionary normative consequences. Thus, we all have some kind of *motivation* (which is not the same thing as a *justification*) to avoid theories which have such consequences, particularly when the revision they imply is *radical*.

2. Two theses regarding the link between phenomenality and value

I now turn to the idea that there is a link between phenomenality and value. I will present and discuss two theses regarding the relation between phenomenality and value. Let us start with the weaker of the two – an attractive, widely accepted thesis I call “*Phenomenal Value*”.

***Phenomenal Value*:** The instantiation of some standard phenomenal properties grounds (or would ground) intrinsic value.

Phenomenal Value states that there are at least *some* phenomenal properties such that their instantiation fully grounds (or would fully ground) intrinsic value. It specifies that these properties must be *standard*. Here, by “standard”, I mean that these properties are *of a kind that phenomenal realists take to be commonly instantiated*. This is necessary in order to make it clear that the thesis would not be made true by some rare and complex value-making phenomenal properties. This thesis is compatible with there being only *one* value-making property (say, phenomenal pain), as soon as it is a standard one.

Phenomenal Value is compatible with the idea that, even if some phenomenal states *fully ground* some intrinsic value, the *total value* of the situations in which they take place can sometimes only be *partially* grounded in these phenomenal states – for example (but not only)

if the situations include some non-phenomenal states also (or complex states including non-phenomenal states) fully grounding value. Importantly, the thesis is formulated in such a way (“grounds (or would ground)”) that *even an illusionist could accept it*. Indeed, even someone who believes that phenomenal properties are *never instantiated* in reality could also believe that, *were some of them to be instantiated*, they would ground value.

This thesis, I think, is very widely accepted. In fact, many philosophers who believe that there is a link between phenomenal consciousness and value have gone further than that. They have not only stated that *some* phenomenal states ground value, but also that *only* phenomenal states could do it, thus endorsing a stronger thesis, which I call *Exclusive Phenomenal Value*.

Exclusive Phenomenal Value: The instantiation of some standard phenomenal properties grounds (or would ground) intrinsic value **and** there is no non-phenomenal property such that its instantiation grounds intrinsic value.

This thesis conjoins *Phenomenal Value* with the restrictive thesis according to which no instantiation of a non-phenomenal property fully grounds intrinsic value – note that this is merely supposed to hold for *instantiated* non-phenomenal properties.⁴ A very rough approximation of this thesis is that “nothing has value without consciousness”. Note that *Exclusive Phenomenal Value* does not prevent non-phenomenal properties from *partially grounding* value. It is therefore compatible with the thesis that value does *not* supervene on the total set of phenomenal states of a situation, and it can accommodate the intuition that, for example, life in an experience machine is not as valuable as life in a “real” world, even if they

⁴ This thesis might also come in a stronger flavor (***Strong Exclusive Phenomenal Value***), in which *Phenomenal Value* is conjoined with the following proposition: “there is no non-phenomenal property such that its instantiation grounds (or would ground) intrinsic value”. While *Exclusive Phenomenal Value* states that there are no *actually instantiated* non-phenomenal properties which grounds value, *Strong Exclusive Phenomenal Value* goes further and asserts that, even in the realm of uninstantiated properties, there is no non-phenomenal value-making property. I will set aside here such subtleties and focus merely on *Exclusive Phenomenal Value*.

are phenomenally identical. Note also that, as I think of it, the second clause of *Exclusive Phenomenal Value* should be read in a way that does *not* prevent non-phenomenal properties to ground value *if they do it by being identical with phenomenal properties, or by grounding phenomenal properties*.

Exclusive Phenomenal Value, compared to *Phenomenal Value*, is quite a strong thesis – though many philosophers have (at least implicitly) endorsed it. Again, it is formulated in such a way that, in principle, it is not incompatible with illusionism.

3. Support for the link between phenomenality and value

What support is there in favor of these two theses? First, it is worth considering that they are entailed by a great many philosophical views. Numerous views state that nothing in reality grounds intrinsic value but some phenomenal states (or larger states partially composed of phenomenal states), such as, notably, phenomenal pain and phenomenal pleasure. This entails *Exclusive Phenomenal Value*. This is notably the case of hedonism⁵, according to which nothing is good or bad but felt pleasure and pain – a view which has found prominent defenders in Ancient times (the Cyrenaics and the Epicureans) as well as in the 18th and 19th century (La Mettrie, Bentham, Mill) and (with some complexities and qualifications) more recently (Crisp, 2006, p. 103-109; Sprigge, 1988, Chapter 5; Tännsjö, 1998, p. 84; for an overview, see Moore, 2013). Besides, even philosophers who do not embrace hedonism have often claimed that nothing can ground intrinsic value but phenomenal states (van der Deijl, 2019; Seager, 2001;

⁵ Of course, this is only true of hedonism if hedonism is conjoined with a *phenomenalist* understanding of pleasure and pain (according to which pleasure and pain are particular phenomenal feelings) and not, for example, to an *attitudinal* understanding of these mental states, according to which pleasure and pain are intentional states (Feldman, 2002). Traditional hedonistic theories were unclear on whether they take pleasure and pain to be constituted by certain *feelings*, or by *attitudes* towards feelings, although I think it is fair to say that most of them were implicitly phenomenalist. This, at least, is very likely to be true for 18th and 19th century hedonists such as Bentham and Mill (as claimed by Moore, 2013) but not of Sidgwick, who is arguably one of the first modern thinkers who developed an attitudinal view of pleasure (Sidgwick, 1907).

Kriegel, forthcoming, 2018, Chapter 9; Brentano, 1952). Some other philosophers state that phenomenal consciousness is *valuable for itself*, which entails at least *Phenomenal Value* (Glannon, 2016; Nagel, 1979; for a discussion, see A. Lee, 2018). Others assert that most (if not all) of the value of our own lives depends on us being phenomenally conscious (Siewert, 1998, 2014), which comes close to endorsing *Exclusive Phenomenal Value*.

Some philosophers have recently tried to downplay the importance of phenomenal consciousness when it comes to value (Carruthers, 1999, 2004; G. Lee, 2014; Levy, 2014). However, it is worth noting that their goal is often to insist that some non-phenomenal mental states can have value, which entails a rejection of *Exclusive Phenomenal Value*, but not necessarily of *Phenomenal Value*. Not all of them explicitly question the idea, for example, that the instantiation of certain phenomenal properties in itself grounds value. This idea seems in fact to be accepted in (Levy & Shepherd, forthcoming)) and entails *Phenomenal Value* (G. Lee, 2014 comes close to denying *Phenomenal Value*, but he focuses on the epistemic value of consciousness and does not say much about the other kinds of intrinsic value it might ground).

In conclusion, numerous views endorsed by philosophers throughout history entail *Exclusive Phenomenal Value*. A number of them reject *Exclusive Phenomenal Value* but remain compatible with *Phenomenal Value* (which they often seem to implicitly accept); and it is difficult to find a view which explicitly denies *Phenomenal Value*.

Now, what about *arguments* in favor of such theses? While many such arguments could be devised, I will simply present one for each thesis. Let us start with *Phenomenal Value*. Given that it is enough to show that there is *one* standard value-making phenomenal property to prove the thesis, I think the strongest arguments focus on the case of the best candidates to this status: valenced phenomenal properties, and notably phenomenal pain. Let us consider the following situation.

Pure Suffering: Let us imagine that a man named Piotr, a genuinely conscious being endowed with phenomenal states, suffers from a rare disease. Every night, during sleep, he wakes up and, for one hour, feels the most excruciating pain one can imagine. The pain is so intense that it leaves him unable to move or to talk (or to have any observable behavior). For one hour, his life is entirely filled by nothing but the most awful pain ever felt (Consider the worst kind of dental pain multiplied a hundred times over): he undergoes *Pure Suffering*. He then goes back to sleep. In the morning, he systematically forgets about everything that has happened to him during the night: these nocturnal episodes of pain do not leave any memory, conscious or unconscious, and do not change his psychological dispositions.

I take it to be extremely intuitive that *Pure Suffering* has negative value. I also take it to be very intuitive that it has negative value *in virtue of the pain felt by Piotr* (the thought experiment eliminates some – but not all – of the non-phenomenal candidates for the grounding of negative value: trauma, negative memory, aversive behavior, etc.). Therefore, this thought experiment makes it intuitive that *phenomenal pain grounds negative value*, which entails *Phenomenal Value*.

It is important to note that, *prima facie*, even an illusionist can accept this argument for *Phenomenal Value*. After all, even if consciousness does not exist, we can still conceive of it (at least, that is what I presupposed). Which means that *Pure Suffering* is conceivable. Once we conceive of this situation, it elicits intuitive judgments of value – and I take it that almost everyone, illusionists included, will intuit (at least at first glance!) that this situation has negative value in virtue of how Piotr feels.⁶ True, an illusionist thinks that *Pure Suffering* is at

⁶ I suppose that, in order to have the intuition, it is enough for the situation to be *conceivable* – it does not have to be *possible*. As William Seager (who goes even further) puts it: “[Q]uestions of value assignment are independent of the kind of possibility which the thought experiment depends upon or exploits. We can assign, if only hypothetically, value to things that are physically impossible or even metaphysically impossible (perhaps even to things that are logically impossible – I might wish that arithmetic had not “turned out to be” incomplete for example)” (Seager, 2001, p. 3). I will come back later to the problem of ascribing value to *impossible* situations.

least nomologically impossible, but that in itself does not prevent her from recognizing that this situation would be bad if it were actual. Similarly, I can recognize that dragons “are” frightening or that the Garden of Eden “is” a delightful place, even if I think that they are nomologically impossible – as long as I recognize that, *if dragons, or the Garden of Eden, were real*, they would be frightening/delightful.

Aside from this direct argument, it is important to note that *Phenomenal Value* arguably plays a central role in justifying some of our deep normative beliefs, themselves crucial in our practices. Consider this: first, we believe that *conscious states* are often much more normatively important than *non-conscious states*. This is arguably why we care a lot about conscious pain, say, but not about unconscious pain. For example, we use analgesics in surgery to suppress conscious pain, and we do not care much about potential *unconscious, subliminal pain* happening somewhere in the brain of the patient during anesthesia. In other words, we endorse the following thesis, which I call “*Consciousness Contrast*”: allegedly phenomenal states have an importantly different value from their non-allegedly-phenomenal counterparts.⁷

Second, we usually believe that *sentient creatures* (who have allegedly phenomenal pain or pleasure) have more importance (or more dignity, or more rights) than non-sentient creatures. This is arguably why we care (at least a little bit) about what happens to creatures we judge to be sentient (dogs, cows, pigs, birds, etc.) while we do not care – except for instrumental reasons – about creatures and things we judge to be non-sentient (bacteria, trees, rocks, computers, etc.). Debating whether or not some creatures are sentient (particularly for “limit cases” such as fish, lobsters, bees, etc.) usually directly affects our opinions on whether or not we should care about

⁷ I use the “allegedly phenomenal” locution here because I want to describe this normative thesis (as well as the next one, *Sentience Contrast*) in a way that remains neutral between illusionism and realism. *Nota bene*: (1) it is only for the phenomenal realist that allegedly phenomenal pain is the same as phenomenal pain. The illusionists think it is only *quasi-phenomenal pain*. (2) Non-allegedly-phenomenal pain is *not* zombie pain nor quasi-phenomenal pain, as it is arguably functionally different from allegedly phenomenal pain. Non-allegedly-phenomenal pain is the form of “unconscious pain” occurring in our world that even phenomenal realists will accept as not being phenomenally conscious – for example, subliminal pain. Some people would argue that such a state does not fully deserve the name “pain” – but nothing substantive here depends on this semantic decision.

them. Generally speaking, we seem to consider that allegedly phenomenal consciousness is crucial for the status of a creature as a genuine moral patient (Levy & Shepherd, forthcoming; Seager, 2001; Sebo, 2017; Shepherd, 2016; Singer, 2011). In other words, we endorse the following thesis, which I call “*Sentience Contrast*”: creatures that enter allegedly phenomenal pain and pleasure (“sentient creatures”) importantly matter more than the creatures that do not (“non-sentient creatures”).⁸ *Consciousness Contrast* and *Sentience Contrast* are two very intuitive normative theses, which are part of the normative consensus, and are arguably crucial for many of our practices.

I think the conjunction of *Phenomenal Value* with phenomenal realism is the thesis that most people would naturally appeal to in order to justify these two theses. The reason why allegedly phenomenal states have a different value from their non-allegedly-phenomenal counterparts is that they are *phenomenal*, and that phenomenal states (or at least some of them) ground intrinsic value. And it is this extra value had by such states that gives in turn more importance (or dignity, or rights) to the creatures able to enter into such states – sentient creatures.

Now that we have examined the support for *Phenomenal Value*, as well as its role in justifying some central normative beliefs, let us turn to *Exclusive Phenomenal Value*. What are the arguments in favor of this thesis? I think that the best arguments appeal to thought experiments featuring worlds devoid of consciousness, which aim to elicit the intuition that such worlds have no value. Of course, there is disagreement about *which worlds* devoid of consciousness are possible/conceivable. Some people think that a possible/conceivable world devoid of consciousness is also a world devoid of creatures with our physico-functional properties, or maybe even a world devoid of a certain kind of functional/informational complexity. It seems plausible that, if someone has such an *austere* conception of a world

⁸ Again, this normative thesis as such is neutral between illusionism and realism.

without consciousness, intuitions that such worlds are valueless are easier to generate. On the other hand, people who believe in the conceivability/possibility of zombies will have a “richer” conception of what worlds devoid of consciousness can be, which hypothetically makes intuitions that such worlds are valueless harder to get. So, the most challenging (and therefore, the most robust) way to argue in favor of *Exclusive Phenomenal Value* is through thought experiments involving *zombie worlds*: worlds identical to ours from the physico-functional point of view, but devoid of consciousness.⁹ Given that I presupposed earlier that zombies are conceivable, it seems in any case best to take that route. Let us consider the following thought experiment.

Zombie world: *Zombie world* is exactly identical to our world from the physico-functional point of view. Moreover, *zombie world* contains no phenomenal experiences. No one in *Zombie world* enters phenomenal pain, phenomenal pleasure, or any other phenomenal states: there is nothing it is like to be any of its creatures. *Zombies*, however, have behaviors, and an internal physico-functional composition identical to ours: when you hurt them, they shout, they cry, the nociceptive zones in their brains are activated exactly like ours, etc.

Does *Zombie world* have value? My own intuitions are not particularly clear at this stage – and I suppose that readers will be divided here. Some people will intuit that such a world has exactly the same value as a world populated with genuinely conscious beings. Others will intuit that it does not, but that it still has *some* value (for example, grounded in some non-phenomenal features that ground value in our own world: informational complexity, cosmic beauty/harmony, the existence of true representations, the existence of functionally conceived

⁹ To the best of my knowledge, the first zombie argument in favor of a link between phenomenality and value has been put forward by Charles Siewert (Siewert, 1998), in order to support the view that phenomenal consciousness is essential to our lives having value. Note that Siewert’s thought experiment is a *zombification* thought experiment, in which a conscious being *becomes* a zombie.

desires and attitudes, etc.). However, I suspect that a number of readers would tend to answer that *Zombie world* has no value; that its existence is entirely indifferent. This intuition is the one needed to support *Exclusive Phenomenal Value*. Note that an *illusionist* (who believes that we live in *Zombie world*) could in principle have any of these intuitions, and that she could therefore end up accepting this argument for *Exclusive Phenomenal Value*.¹⁰

4. The normative challenge for illusionist views of consciousness

That there is a link between phenomenality and value – the one expressed by *Phenomenal Value*, or the stronger link expressed by *Exclusive Phenomenal Value* – is, I think, quite an attractive position. As I showed, it is entailed by a great variety of influential philosophical views, it is directly supported by arguments, and it plays (in its weak version, when conjoined with phenomenal realism) an important role in justifying some of our central normative beliefs and practices.

Because this position is attractive, it creates a challenge for illusionists: the normative challenge. Illusionists must decide whether or not they accept that there is such a link between phenomenality and value, and say where they stand regarding these two theses. Depending on their answer, they might be led or not to revisionary normative consequences. *If* they end up being committed to such consequences, the challenge for them is to explain what these normative consequences exactly are, see if they can be limited in scope, and eventually show how we can live with them. *If* they avoid being committed to revisionary normative consequences (which, I will show, probably requires them to reject *Phenomenal Value*), the

¹⁰ One might note that, by formulating this thought experiment in terms of “zombies”, we might tend to presuppose (implicitly) the falsity of illusionism and the truth of phenomenal realism. This could in turn pump our normative intuitions in some particular directions (possibly problematic from the point of view of an illusionist). For more on the difficulties we might face when we try to get rid of our *realist* intuitions, as well as the effect it might have on our intuitions regarding what has value, see section 6.3, item iv.

challenge for them is to explain why the idea that there is such a link between phenomenality and value seems true even though it is false. Moreover, rejecting *Phenomenal Value* does not entirely guarantee illusionists against *all* revisionary normative consequences; they still have to make sure that their view of the non-phenomenal ground of value salvages our intuitive view of what has value. If it does not, then again they have to clarify which exact revisionary normative consequences they accept, and show how we can live with them.

Let us now examine the options an illusionist can take regarding the thesis that there is a link between phenomenality and value. There are three main options: (1) the nihilist option, (2) the moderate option, (3) the conservative option.

(1) *The nihilist option. Accept Exclusive Phenomenal Value.* If an illusionist accepts *Exclusive Phenomenal Value*, she is naturally committed to *radical revisionary normative consequences*. Indeed, she has to say that there are no instantiated properties in the world that ground intrinsic value: therefore, that no situation in the world has intrinsic value. This position is a *nihilist* one. Nihilism creates all kinds of difficulties (including “pragmatic” ones: if nihilism is true, it means that it is false that we *should* believe so, which maybe means that we cannot believe it (Streumer, 2013)) but one cannot assume that it is false – and I will not here. After all, even if most philosophers judge it to be unattractive, it has been defended independently of issues regarding phenomenal consciousness (Mackie, 1977). However, we can also admit that the view that nothing has value is radically revisionary, and that it will not be easy to answer the normative challenge if one takes that option: that is, that it will not be easy to *learn to live with it*.

(2) *The moderate option. Reject Exclusive Phenomenal Value, accept Phenomenal Value.* By denying *Exclusive Phenomenal Value*, the illusionist avoids the

commitment to nihilism. At the same time, she might still accept *Phenomenal Value* – the weaker, more widely accepted, most intuitively supported thesis – which means that she grants that the instantiation of *some* phenomenal properties (maybe at least phenomenal pain and pleasure?) *would ground* intrinsic value. In that case, I will show that it is extremely difficult to avoid at least *some* revisionary normative consequences. The illusionist who chooses the moderate option then has to face two tasks. First, she must explain precisely which revisionary consequences she accepts – ideally, showing that these consequences can be limited. Second, she has to justify her denial of *Exclusive Phenomenal Value*.

(3) *The conservative option. Reject Phenomenal Value.* The illusionist who rejects *Phenomenal Value* (and, therefore, *Exclusive Phenomenal Value*) has to reject a very intuitive and almost universally accepted thesis, which is a costly move. One advantage is that she is probably in a better position to avoid revisionary normative consequences. However, even she is not entirely guaranteed against such consequences: she still has to give a view of the non-phenomenal ground of value, which makes it clear that such consequences can be avoided (or at least limited). Moreover, she has to justify her denial of *Phenomenal Value* – which is no trivial task.

This is the structure of the normative challenge. Illusionists must take one of these three options, and all of them create difficulties. These difficulties are quite obvious in the case of the *nihilist option*. The normative challenge facing the illusionist who takes this route is not unique to illusionism: it is “simply” the more general problem of nihilism. Given that nihilism has been widely discussed by philosophers for at least a century, I will not explore this option

any further here – which does not mean I do not take it to be worth discussing. I will focus on the two other options: the moderate option, and the conservative option.

5. Answering the normative challenge: the moderate option

5.1. The denial of *Exclusive Phenomenal Value*

Start with the moderate option: denying *Exclusive Phenomenal Value* while accepting *Phenomenal Value*. The “moderate” illusionist first has to justify her denial of *Exclusive Phenomenal Value*. *Exclusive Phenomenal Value* is supposed to be supported by the *Zombie world* thought experiment, which arguably elicits the intuition that zombie worlds are valueless. This means that the “moderate” illusionist must either: (1) try to elicit different intuitions about such a case, or (2) give a reason not to trust intuitions that zombie worlds are valueless. I think that strategies of the second kind usually also allow one to discard intuitions supporting *Phenomenal Value* itself, which is why I postpone their presentation to the next section. As for strategies of the first kind, I suspect they have respectable chances of success. Indeed, I do not think that the intuition that zombie worlds are valueless is overwhelming. One can point at various aspects of a zombie world which seem intuitively likely to ground value. For example, zombie worlds are *diverse and complex*. Arguably, they contain as much “natural” beauty (corresponding to the complex structures of the physical world, or to the harmonious complexities of mathematics) as a world with phenomenal consciousness. Moreover, zombie worlds feature *complex organized systems*. Some of these – zombies – are “cognitive” systems, able to *want*, to *believe*, to *hope* (although in a purely functional, non-phenomenal way). I find it not entirely intuitively implausible that this could ground *some* value. One strategy to elicit the intuition that zombie worlds have *some* value (partly inspired by the « Phenumb » thought

experiment presented by Peter Carruthers Carruthers, 1999)¹¹ might be to focus our attention on the zombie twin of a human being and try to imagine that something really awful happens to this creature. Let us consider the zombie version of the *Pure Suffering* thought experiment.

Pure Zombie Suffering: Zombie Piotr, the zombie twin of Piotr (from *Pure Suffering*), suffers from a rare disease. Every night, during sleep, he wakes up and, for one hour, enters into a state that is physically and functionally indistinguishable from the state of Piotr when Piotr feels the most excruciating pain one can imagine. For this one hour, Zombie Piotr is unable to move or to talk (or to have any observable behavior), in virtue of the functional properties of the state he is in. Even though he does not have any phenomenal experience, all of his functionally conceived mental states – desires, judgments, etc. – are similar to the ones had by Piotr when Piotr suffer: he undergoes *Pure Zombie Suffering*. Zombie Piotr then goes back to sleep. In the morning, he systematically forgets about everything that has happened to him during the night; these nocturnal episodes of zombie pain do not leave any memory, conscious or unconscious, and do not change his psychological dispositions.

During *Pure Zombie Suffering*, Zombie Piotr is in zombie pain. True, he does not feel phenomenal pain, but he still keeps judging (functionally) that everything is awful, he keeps having the (functional) ardent desire for his current state to stop, etc. Some of us, I think, would have some kind of empathy for this creature, and ascribe some kind of negative value to his state. Note that one does not need here to intuit that *Pure Zombie Suffering* is *as bad* as *Pure Suffering*: all that is needed is the intuition that *Pure Zombie Suffering* is not entirely indifferent and that it has *some kind* of negative value, in order to reject the claim that zombie worlds are

¹¹ Phenumb is a conscious agent gifted with with conscious desires and conscious beliefs, but who does not feel any conscious frustration or sadness whenever his desires are frustrated. Carruthers argues that, when Phenumb is frustrated or sad, he is an appropriate object of sympathy even if his sadness/frustration are not conscious.

valueless. Whether or not such intuitions about the badness of zombie pain are resistant upon reflection, and whether it can lead to satisfying theories regarding the non-phenomenal ground of value, will be examined later.

5.2. The moderate option and its revisionary normative consequences

Let us assume that the moderate illusionist has successfully rejected *Exclusive Phenomenal Value*. Given that she accepts *Phenomenal Value*, where does she stand when it comes to revisionary normative consequences? I think that she will very probably have to accept at least *some* revisionary normative consequences. Indeed, the illusionist who accepts *Phenomenal Value* thinks that the instantiation of some standard phenomenal properties would ground intrinsic value. As an illusionist, she also thinks that human beings are intuitively phenomenal realists – as they are under the illusion of phenomenality. So, the moderate illusionist accepts that many real situations do not feature the phenomenal properties that they are intuitively thought to feature, and she accepts that these phenomenal properties *would ground* intrinsic value. At this point, the only way for the illusionist to avoid the conclusion that these real situations *do not have as much value* as they would have if, as people intuitively think, they had phenomenal properties, would be to say this: in all these situations, the exact same intrinsic value that *would be grounded* by phenomenal properties *is always already grounded* by some *non-phenomenal properties*. But this would require an *amazing coincidence*, as this would mean that, in the case of the *phenomenal version* of such situations (the one which realists take to be actual), the same intrinsic value always and systematically would have *two distinct full grounds*.¹² Supposing such an amazing coincidence seems exceedingly *ad hoc*. Therefore,

¹² Another more attractive, less *ad hoc* option for the illusionist is of course to say that there are not *two full grounds* of value in all of these situations, but only *one full ground* – which is *not* phenomenal. But this option requires one to deny *Phenomenal Value*, and it will therefore be examined later.

embracing illusionism and *Phenomenal Value* very probably leads to revisionary normative consequences.

What are exactly these revisionary normative consequences? This depends on the exact view one embraces regarding which phenomenal properties *would ground value*. If one grants, for example, that phenomenal pain and phenomenal pleasure would ground value (negative or positive), the natural consequence will be that allegedly phenomenal pain and pleasure *are simply not as good and bad* as what they would be if the realist were right (which does not mean they have no value at all). This arguably extends to the case of other putative value-making phenomenal properties.

This can very well appear to be an uncomfortable consequence for illusionists: do they really want to say that their theory implies that pain is not as bad as what we originally thought? The concrete normative consequences are potentially quite dramatic: it could for example imply that (say) torture or factory farming (and other things that we usually judge to be *bad* at least partly because they involve *pain*) are not as bad as we thought.¹³ This could in turn have consequences regarding what we *should do* regarding these things (providing some assumptions about the link between values and duties).

On the other hand, the moderate illusionist can say, for example, that allegedly phenomenal pain (which, for her, is quasi-phenomenal pain – in what immediately follows, I will simply write “pain”), *is* bad, for reasons that have nothing to do with its (in any case nonexistent) phenomenality (though, barring the *amazing coincidence* I mentioned earlier, she has to accept that it is not *as bad* as it would be if the realist were right). However, she then faces two

¹³ On a side note: if someone holds the view that *nothing grounds value in the world but allegedly phenomenal states*, conjoined with *Phenomenal Value*, and then comes to embrace illusionism, she is certainly committed to the idea that allegedly phenomenal states are not as good/bad as what she thought. However, depending on how she then determines the value of the states she thinks are actual (quasi-phenomenal states), she may simply “rescale” all of the values of real situations, so that the *ordering* of the value of all situations remains similar to what she believed before embracing illusionism. She would conclude that things are not as good or as bad as what she thought, but that all comparative relations of value (better/worse) still hold exactly as before. Such an option is closed if the subject holds the view that at least one actual thing outside of allegedly phenomenal states grounds value. Thanks to Johannes Martens for pressing this point.

challenges. First, she has to *give reasons* to think that pain is bad; second, she has to ultimately give a theory of *what makes* pain bad (the same thing would happen with other states, but here I focus on pain).¹⁴ This is no trivial task for an illusionist.

First, it is no trivial task for the illusionist to *justify* the thesis that pain is bad, because the illusionist cannot appeal to the fact that the badness of pain is immediately manifest through introspection – at least not in the same way the realist can. True, the illusionist can justify the badness of pain by the fact that we usually *judge* that pain is bad, and even that we usually *judge immediately* (i.e. intuit) that pain is bad when it happens to us or to others.¹⁵ However, the illusionist cannot say – contrary to *some* realists – that the nature of pain is entirely and immediately revealed to us. The idea that introspection reveals the entire nature of phenomenal states is one of the tenets of the traditional realist epistemology of phenomenality. The illusionist takes introspection to be illusory, and therefore cannot appeal to such a principle: our pain, which is really quasi-phenomenal pain, is not revealed to us. However, without such epistemology (or something in the vicinity), why should we particularly trust our judgments, or even our immediate judgments (i.e. intuitions), about the *badness* of pain – or of anything? Human history, after all, is full of widespread normative judgments and intuitions that we now take to be completely false and misguided. Note that, on the other hand, realists who endorse something akin to the traditional epistemology of phenomenal states are often inclined to think

¹⁴ I differentiate these two steps because I think they can be separated. For example, let us consider the statement: “Mary’s car is yellow”. I can *justify* this statement by appealing, say, to perception or to testimony; I can then give an *explanation* of what makes this statement true (appealing to the reflectance profile of the surface of Mary’s car). These two processes seem quite different, although they can eventually be linked.

¹⁵ This seems to be the strategy suggested by Luke Muehlhauser when he states that illusionism regarding consciousness might have no consequences when it comes to the value of pain: “What are the implications of illusionism for my intuitions about moral patienthood? In one sense, there might not be any. After all, my intuitions about (e.g.) the badness of conscious pain and the goodness of conscious pleasure were never dependent on the “reality” of specific features of consciousness that the illusionist thinks are illusory. Rather, my moral intuitions work more like the example I gave earlier: I sprain my ankle while playing soccer, don’t notice it for 5 seconds, and then feel a “rush of pain” suddenly “flood” my conscious experience, and I think ‘Gosh, well, *whatever this is*, I sure hope nothing like it happens to fish!’ And then I reflect on what was happening *prior* to my conscious experience of the pain, and I think ‘But if *that* is all that happens when a fish is physically injured, then I’m not sure I care’” (Muehlhauser, 2017). Muehlhauser actually rejects this interpretation of his position (as this passage simply describes one of his moral intuitions, but not his *ultima facie* moral judgment), but I nevertheless think that it would constitute an interesting position, attractive at least to some illusionists.

that statements of value concerning phenomenal states (for example, concerning the badness of pain or the goodness of pleasure) are uniquely justified compared to other statements of value. Seager, for example, stresses the contrast between our apprehension of the value of phenomenal states and of other states of affairs. “The general problem here is that although it is easy to see why states of consciousness are intrinsically valuable since they have a subjective component that, so to speak, directly reveals their value [...] it is impossible to see why any non-conscious state should have intrinsic value [...] One can, of course, *assert* that diversity, say, is intrinsically valuable, but unless diversity is tied to some conscious being’s *appreciation* of diversity, which it almost invariably is [...] the assertion seems completely empty and devoid of support” (Seager, 2001, p. 4).¹⁶

Of course, many phenomenal realists (especially the ones who are also materialists) do not endorse such a strong epistemology of phenomenality: does that mean that they, too, are in difficulty when it comes to justifying the badness of pain? It might be, one could answer, that we are just setting the bar too high here: in order to justify our belief that pain is bad, we do not really need to say that this badness is grasped through a process of introspection and that introspection *reveals* phenomenal states. We simply need to say that badness is grasped through introspection and that introspection is a *reliable* (though fallible and imperfect) way of forming beliefs about such states. But here one can see that there is a difference between the realist and the illusionist’s position. Phenomenal realists with a “weak” epistemology can say that introspection is not *revelatory*, but constitutes a *reliable* way of acquiring information about the nature of phenomenal states, so that we would be justified in trusting our introspective verdicts regarding the badness of pain. But the illusionist cannot as easily appeal to this kind of weaker thesis, as she claims that introspection systematically generates incorrect representations about the nature of our internal states. This casts a serious doubt on the

¹⁶ For the idea that the badness/goodness of pain/pleasure are uniquely justified, see also (Sinhbabu, manuscript).

reliability of introspective judgments – including the judgments according to which, for example, pain (the pain we *really* have: quasi-phenomenal pain) is bad.

At this point, I suspect that the best strategy is for illusionists to appeal to other, non-introspective ways to justify the badness of pain. In order to do the job, however, illusionists have to make sure that no hidden appeal to introspection is made. We might want to say, for example, that we know that pain is bad through the testimony of others, but it is likely that others' claims must themselves be justified by introspection. Alternatively, one might want to say that we know that quasi-phenomenal pain is bad *a priori*, simply by reflecting on it: but it is unclear how reflection on a physical-functional state could deliver such intuition. Moreover, this would arguably imply that, as long as we do not have a correct physical-functional theory of quasi-phenomenal pain, we cannot justify that it is bad, which seems extremely unsatisfying. Finally, one might suggest that we know directly that pain is bad on the basis of *empathy*: perceiving with empathy that someone is in pain gives us a direct grasp of the badness of their pain. It is notable to this way of knowing the badness of pain that I appealed earlier when I tried to elicit the intuition that *Pure Zombie Suffering* is bad: we can maybe feel empathy towards Zombie Piotr when he undergoes zombie pain, which makes us *know* that his pain is bad. It remains to be seen if such a “second-person” epistemology of the value of mental states can be worked out in details (after all, we can be made to feel empathy for many things, including inanimate objects), and if it can really solve the illusionist's difficulty.

Another difficulty is that it is no trivial task for the illusionist to explain *what makes pain bad* – what grounds its badness. She cannot appeal to its intrinsic phenomenal quality, given that she thinks pain is not phenomenal. Appealing to some kind of rationalistic moral principle (for example, a Kantian principle) will probably give the verdict that *some* situations in which someone is in pain are *bad* (for example, situations in which someone is treated as a mere means), but it will not be the case of all of them – it is unclear, for example, why Kantian

principles would imply that *Pure Zombie Suffering* is bad. Looking at quasi-phenomenal pain from a purely physiological point of view (as a brain state), on the other hand, is unlikely to give us a satisfying account of its badness: nothing in the mere description of a brain state seems fit to explain the presence of value. The best option here for the illusionist would probably be to draw inspiration from desire-satisfaction views of well-being (Brandt, 1979; Heathwood, 2006) and/or from attitudinal theories of valenced states (Feldman, 2002), and to say that pain is bad (even if it is not phenomenal) because it constitutively includes the frustration of a desire, or the having of a certain negative attitude of *dislike*. After all, when I am in pain, there is something awful which is that I *want it to stop* (and that my desire is frustrated); alternatively, one could insist on the fact that what is bad is that I *dislike* my pain. This *frustration* or this *dislike* are what makes pain a *harm*, which in turn grounds its negative value. This might be the most promising lead to an account of what makes pain bad.

I said earlier that the conjunction of *Phenomenal Value* with phenomenal realism was the natural justification of two central normative theses: *Consciousness Contrast* and *Sentience Contrast*. Can the illusionist who takes the moderate option conserve these two theses?

The illusionist who accepts *Phenomenal Value*, is forced (barring the *amazing coincidence*) to admit that allegedly phenomenal states do not ground as much value as they would have if the realist were right: which means that the contrast between allegedly phenomenal states and non-allegedly-phenomenal states cannot be as stark as if the realist were right. That such contrast has to be weakened is already an important revisionary normative consequence. If the contrast of value between the states is what justifies the contrast of importance (or dignity, or rights) between the creatures, this means *Sentience Contrast* also has to be weakened.

I said “weakened”, because illusionists can still say that there is *some kind of difference in value* between allegedly phenomenal and non-allegedly-phenomenal states (which can lead to a special status for sentient creatures). But again, the illusionist will have to first *justify* that

such a difference holds – and, again, this will be difficult, as illusionists cannot appeal to the same kind of epistemology of phenomenality as the realist. Second, the illusionist has to *explain* what grounds such difference in value. But this might be particularly tricky: indeed, if the illusionist has previously accounted for the badness of pain, say, in terms of its inclusion of a frustration of a desire (or of a negative attitude of dislike), it will not be easy to explain why there is a stark contrast in value between quasi-phenomenal pain and non quasi-phenomenal pain. After all, are not both frustrations of desires of some kind (or dislikes of some kind)? Peter Carruthers (Carruthers, 2004), who defends a similar view of what grounds the badness of pain (although he is not an illusionist), seems to concur with this idea. He thus emphasizes that the harm (that I interpret here as grounding some negative value) constituted by non-conscious pain (i.e., non-allegedly-phenomenally conscious pain) is not as radically distinct from the harm constituted by conscious pain (i.e., allegedly phenomenally conscious pain) as what is usually thought. But this endangers the justification of *Consciousness Contrast* (and, hence, of *Sentience Contrast*).

One might emphasize the difference between the value of the two kinds of states by appealing to differences in the frustrated desires (or differences in the attitudes of dislike). Maybe it is worse to frustrate a quasi-phenomenal desire than a non quasi-phenomenal desire; or to dislike something in a quasi-phenomenal way than in a non quasi-phenomenal way. However, this in turn seems to stand in need for a justification – why should we believe that it is worse? – and an explanation – what *makes it* worse exactly? At any rate, such a view would need to be elaborated in details.

To close this section: the illusionist who takes the moderate option must very probably face *some* revisionary normative consequences. Some mental states do not have the value they would have if the realist were right, which also leads to a weakening of the contrast between states (allegedly phenomenal vs non-allegedly-phenomenal) and creatures (sentient vs non-sentient).

Moreover, if illusionists wants to *limit* these normative consequences – and maintain, say, that some valenced quasi-phenomenal states ground value, and that *Consciousness Contrast* and *Sentience Contrast* still hold in a weakened version – she has to give a *justification* for such theses, and she has to explain what makes them true. As I tried to show, this is no trivial task.

6. Answering the normative challenge: the conservative option

6.1. The conservative option and its revisionary normative consequences

Let us now consider the conservative option. Illusionists who take this route reject *Phenomenal Value*. That seems to put them in a better position to avoid revisionary normative consequences. Indeed, if they reject *Phenomenal Value*, they deny that the instantiation of phenomenal properties *would ground* value. Hence, there seems to be no particular reason to think that embracing illusionism should change our view regarding *what* has value.

However, two things must be noted. First, rejecting *Phenomenal Value* does not guarantee against *all* the normative revisionary consequences of illusionism (even if it helps). After all, it might be that *no phenomenal states would fully ground value* (i.e. *Phenomenal Value* is false) but that *some phenomenal states would partially ground value* if they were the case, in such a way that illusionism would have different consequences than realism regarding what currently has value.¹⁷ Second, it could be that the illusionist who embraces the conservative option is *right* to reject *Phenomenal Value*, but that the intuitive default position of most people is to embrace *both Phenomenal Value* and phenomenal realism, and to thus ascribe a certain value

¹⁷ Here is an example (rather innocuous): if phenomenal states were the case (as the realist think they are), *beliefs* (that most people have) that we enter into such phenomenal states would be true, which in turn would arguably ground some kind of (epistemic) value. The “same” beliefs do not ground the same value if the illusionist is right. So, that would be at least one (rather modest) revisionary normative consequence that illusionism could have even if *Phenomenal Value* is false (as it does not presuppose that phenomenal states fully ground value).

to *some* allegedly phenomenal states in virtue of them really having a certain phenomenal feel. By rejecting *Phenomenal Value*, the illusionist can then maybe deny that allegedly phenomenal states would have a different value if they were phenomenal (illusionism does not have revisionary normative consequences compared to realism, assuming *Phenomenal Value* is false). However, she cannot guarantee that allegedly phenomenal states have the value that most people intuitively ascribe them – she cannot guarantee that illusionism *conjoined with the rejection of Phenomenal Value* does not have revisionary normative consequences compared to the intuitive default position – realism conjoined with *Phenomenal Value*. So, her overall position might still be revisionary compared to the intuitive starting point.

If the illusionist wants to make sure that major revisionary normative consequences are avoided – for example, if she wants to make sure pain and pleasure are as bad as intuitively thought, or if she wants to justify *Consciousness Contrast* and *Sentience Contrast* in its strong version, she has to provide two things. Like the illusionist who takes the moderate option, she first has to provide a *justification* for the theses that (for example) pain is bad and pleasure is good (and that allegedly phenomenal pain/pleasure are *especially* bad/good when compared to non-allegedly-phenomenal pain/pleasure). She also has to provide an *explanation* of why these theses hold. However, for the very reasons presented in the previous section, illusionists might encounter difficulties here. The justificatory part is made tricky because illusionists reject the realist epistemology of phenomenality. The explanatory part may feature some tension between an account of the value of valenced quasi-phenomenal states in terms of desire satisfaction/frustration or in terms of attitudes (like/dislike) on the one hand, and an account of what makes *Consciousness Contrast* and *Sentience Contrast* true on the other hand. So, rejecting *Phenomenal Value* puts illusionists in a better position to avoid revisionary normative consequences, but they still face some serious challenges if they want to do so.

Moreover, rejecting *Phenomenal Value* presents a difficulty of its own. Indeed, *Phenomenal Value* is an attractive and intuitively well-supported thesis, - arguably, much more than *Exclusive Phenomenal Value*. That means that justifying the rejection of such a thesis is no trivial matter. The illusionist who chooses the conservative option has two possibilities: she can (1) try to change our intuitions regarding *Phenomenal Value*, or she can (2) admit that *Phenomenal Value* is intuitively supported, but then discard this intuition, by giving *reasons* not to trust it. I will first explore the first strategy, and then turn to the second one.

6.2. Changing our intuitions regarding *Phenomenal Value*

Start with the first strategy: how could illusionists make the intuition that supports *Phenomenal Value* disappear? Let us consider the thought experiment I called *Pure Suffering*. We have a strong intuition that this situation is *bad*. This intuition is unlikely to change (and changing it would anyway not help the illusionist with dealing with revisionary consequences). But we also have a strong intuition that this situation is bad *in virtue of what Piotr feels* – that is, in virtue of his phenomenal states. This is what gives support to *Phenomenal Value*. Maybe illusionists can pump our intuition in another direction?

One way to do so would be to claim that, if we think harder, we will understand that what we intuitively find bad in Piotr's pain is *not* something grounded in phenomenality. For example, drawing inspiration from the already mentioned attitudinal theories of valenced states (or from desire-satisfaction theories of well-being), illusionists could claim that what is intuitively bad about Piotr's pain is not the way it *feels* (the pain *quale*). In fact, it is (1) the fact that Piotr *dislikes* the way pain feels, and/or (2) the fact that he desires this feeling to stop (and that this desire is frustrated). After all, the same painful "feeling" would maybe not be *bad* if Piotr liked it (or if he wanted it to go on): if a masochist enjoys her pain and wants it to go on,

it seems doubtful that her pain is really *bad*. But if it is not the phenomenal aspect of Piotr's pain that makes it intuitively bad, but his dislike of it/his desire for it to stop, then intuitive support for *Phenomenal Value* seems to vanish.

Can the illusionist successfully pump our intuition in that direction? Maybe. But here we need to approach things with precautions. Indeed, notice that the idea according to which what is bad in pain is not the pain *quale*, but rather the fact that pain is disliked, or that it generates a frustrated desire, only allows to counter *Phenomenal Value* if we use a concept of *dislike* and a concept of *desire* that make it clear that they do not have to be phenomenal (or grounded in the phenomenal). But I precisely suspect that, when we find intuitive the idea that the badness of pain does not come from its *quale* but from us disliking it/wanting it to stop, we implicitly use a *phenomenal* concept of dislike/desire.

One way to make this problem clear is this: if the phenomenality of pain (including the phenomenality of dislikes, and the phenomenality of frustrated desires) is *not* intuitively what makes it bad, then *Pure Zombie Suffering* should be intuitively *as bad* as *Pure Suffering*. Indeed, Zombie Piotr has the same attitude of dislike (non-phenomenally conceived) towards his pain, and the same (non-phenomenal) desire for it to stop, as conscious Piotr. But I find it personally very hard to have such an intuition: I can intuit *some* negative value in *Zombie Pure Suffering*, but intuitively it also seems that there is an important difference with *Pure Suffering*, which makes this situation simply *less bad*.

There is another other way to make it clear that the intuitive plausibility of such accounts of the badness of pain rely on an implicit acceptance of a *phenomenal* conception of dislikes and frustrated desires. It appeals to imaginary cases of *hedonic inversions* (Kahane, 2009, p. 334; cited in Levy & Shepherd, forthcoming)). Let us imagine that Piotr's friend, Natasha (a conscious being), suffers from a neighboring condition, except that instead of feeling pain for one hour every night, she feels bliss. Let us call this situation *Pure Bliss*. Now, let us suppose

that there is a *third character*, “Inverted Qualia” Andrei, who, each night, enters for one hour in a state that is phenomenally *exactly identical* to *Pure Bliss* (Natasha’s state), but functionally and physically exactly identical to *Pure Suffering and Pure Zombie Suffering* (Piotr and Zombie Piotr’s states). Let us call his situation *Inverted Bliss-Suffering*. I take it to be extremely intuitive that such situation is *not* as bad as *Pure Suffering*: how could it be, given that “Inverted Qualia” Andrei feels nothing but phenomenal bliss, instead of phenomenal pain? But if what makes intuitively pain bad is the fact that we dislike it, or the fact that we desire it to stop, and if such dislike and desire are intuitively not phenomenal (nor grounded in the phenomenal), then we should intuitively admit that *Inverted Bliss-Suffering* (and, of course, *Pure Zombie Suffering*) is exactly as bad as *Pure Suffering*.¹⁸ I find it *intuitively* unacceptable – which is not to say here that I claim that they *really* differ in value.

So, I think that the illusionist will have trouble making our intuitive support for *Phenomenal Value* vanish. But can she discard this intuition – accept that this thesis seems true, but give reasons not to trust this seeming?

6.3. Discarding our intuitions for *Phenomenal Value*

For any intuition, however strong, one can always claim that the proposition supported by the intuition is *false*. However, given that intuitions – particularly intuitions which do not disappear on reflection – are usually seen as providing support in favor of what is intuited,

¹⁸ I presupposed earlier that zombies are conceivable, which notably implies that there is no *a priori* derivation from physical truths to phenomenal truths. The conceivability of *Inverted Bliss-Suffering* actually requires more than that, as it also requires that there is no *a priori* derivation from phenomenal truths to physical truths. So one could technically accept the conceivability of zombies, but not of such inverted *qualia*. However, it seems to me that philosophers who usually accept that zombies are conceivable should also naturally accept the conceivability of *inverted qualia*. Moreover, claiming that there is an *a priori* derivation from phenomenal truths to even simply *some* physical truths would have all kinds of strange consequences: it would make phenomenal idealism (the view that nothing exists but phenomenal experience), contradictory, and *a priori* false. That would also make skepticism about the physical world contradictory and *a priori* false. However, such views do not seem contradictory nor *a priori* false.

denying the truth of an intuited claim leaves us in a difficult situation if we are not also able to provide a *reason* not to trust this particular intuition. Here are a few routes that the illusionist willing to discard the intuition supporting *Phenomenal Value* might take:

i/ *Global skepticism regarding intuitions of values*. One way to go would be simply to refuse to give weight to any of our intuitions concerning value – including the intuition supporting *Phenomenal Value*. Some philosophers might find this stance attractive – maybe, for example, error theorists about value (Mackie, 1977). It could be justified on the basis of a debunking argument: because of the way our intuitions about values have been acquired (e.g. through a process of natural selection which is not sensitive to value), we should not trust them (Joyce, 2006; Street, 2006). However, this move is likely to have even *more* revisionary normative consequences than most other options available to the illusionist (in the most radical case, it leads to nihilism). For that reason, I do not consider it a particularly attractive option in this context.

ii/ *Skepticism regarding intuitions of values about non-phenomenal states*. Another strategy would be to discard some of our intuitions supporting *Phenomenal Value* – notably our intuition that *Pure Suffering* and *Pure Zombie Suffering* have a different value – by claiming that our intuitions regarding the value of some large class of states, such as *non-phenomenal states in general*, are generally not to be trusted. This strategy is suggested (in a different debate) by Peter Carruthers (Carruthers, 1999), who claims that we are bad at properly evaluating non-phenomenal mental states, because we cannot *imagine them in the right way*, which makes us conclude that they must have no significance at all. I find that a strategy directly inspired by Carruthers' view is probably unsatisfying, in that I think that one can very well intuitively ascribe *some* negative value to non-phenomenal pain, while also intuiting that there is an important contrast between its value and the value of phenomenal pain – phenomenal pain seeming considerably *worse* than non-phenomenal pain. But if Carruthers were right, such

intuition should not occur, as the cause of our misjudging that there is a stark difference in value between phenomenal pain and non-phenomenal pain would be our incapacity to intuit the negative value of non-phenomenal pain. There might be other, more subtle ways to justify skepticism regarding intuitions of values about non-phenomenal states. However, in order to be satisfying, they generally cannot state that we wrongly only intuit a difference in value between phenomenal states and non-phenomenal states merely because we wrongly intuit that non-phenomenal states have no value.

iii/ *Skepticism regarding intuitions of values about phenomenal states.* On the other hand, one can defend a form of skepticism about our intuition of values regarding *phenomenal* states: our mistake would not be that we wrongly underestimate the negative value of *Pure Zombie Suffering*, but maybe that we *overestimate* the negative value of *Pure Suffering*. Of course, such skepticism has to be justified independently, on pain of being *ad hoc*. One first way to go would be to say that intuitions concerning the value of *Pure Suffering* (for example) are justified by introspection, but that illusionism takes phenomenal introspection to be systematically misleading, so that we should not trust the intuitions it delivers. One problem with this strategy is that it is in fact unlikely that we grasp the badness of *Pure Suffering* *via* introspection properly speaking. After all, none of us (hopefully) has ever lived *Pure Suffering*, which means that *we do not introspect it!* We *conceive* of it, and we grasp its badness through concept application, through something more akin to an *a priori* process (different from introspection, which seems to be an *a posteriori* process). This *a priori* process allows us to intuit what *would be the value of this situation* if it were the case, and even if it is never the case (the same way in which I can grasp the frightfulness of dragons, or the delightfulness of Eden). One might insist that this seemingly *a priori* process is ultimately justified by what we have learnt *a posteriori* about the badness of our own “normal” pain (that we have introspected!). The problem, however, is that it is unclear that it is really how we learn even about the badness of our own pain: it is doubtful

that we introspect that we are in *pain* in the one hand, and that we are in a *bad state* in the other hand. Otherwise, conceiving of an intrinsically *good pain* (or a bad pleasure) would be easy, as we would have two different introspective representations (one for pain, one for badness) whose link is entirely *a posteriori* and contingent. However, conceiving of an intrinsically bad pain/good pleasure does not seem that easy. So, it seems more plausible that we introspect that we are in pain, where pain is essentially characterized as a state with a certain valence, grounding a certain value.¹⁹ But, for that reason, casting doubt on the reliability of introspection is not enough to reject the badness of introspected pain (that is, it is not enough to deny that, *if such state was real, it would be bad*). In the same way, learning that my perception of a dragon is hallucinatory does not undermine the claim that, if the dragon that I hallucinate were real, it would be frightful, big and scaly, given that *my (hallucinatory) perception characterizes the dragon as such*.

Another way to justify skepticism about the value of phenomenal states is to claim that phenomenal consciousness is *impossible* – and then discard intuitions of values regarding *impossible* cases. I supposed earlier that we have trustworthy intuitions of values about all conceivable cases. But this may seem far-fetched, and one could wonder why we should give any kind of weight to our intuitions regarding the value of impossible (though conceivable) situations. Indeed, it is doubtful that we should take seriously theses that state that an impossible situation *has* a certain value. For example, we could say that “impossible worlds” really are just a way of talking, so that statements about impossible worlds do not have a truth-value (or maybe they are all trivially true, or false). We can then discard our intuitions about impossible situations having value, because the proposition intuited does not have a (non-trivial) truth-value. If the illusionist combines this idea with the claim that *phenomenal consciousness* is not

¹⁹ For reasons of space, I cannot discuss it in details, but I want to point out that this view is not endangered by the existence of medical conditions such as *pain asymbolia*, as here I use “pain” to refer to pain *as normal subjects introspect it*.

only inexistent, but also *impossible*, she can conclude that we should discard all of our intuitions regarding the value of phenomenal states (including our intuitions supporting *Phenomenal Value*).

Illusionists willing to take that route still have to fulfill one task: they have to give reasons to think that phenomenal consciousness is impossible. But phenomenal consciousness seems both possible and actual – in fact, almost everyone in history *believed* it to be actual. Giving arguments for its metaphysical impossibility will be no trivial task and will constitute a genuine theoretical cost.²⁰

iv/ *Skepticism about our value intuitions given our realist intuitions*. One other strategy relies on the following idea. According to illusionists, we are all instinctive phenomenal realists, as we are all under the illusion of phenomenality. One can claim that this illusion is so strong that most people (even convinced illusionists), in most contexts, cannot help but systematically fall back into the illusion of phenomenality and believe, deep down, that phenomenality is real. This “gut” belief would then pervade our thought experiments, and make us (falsely) believe that *Pure Zombie Suffering* simply is *not* what *we would undergo* if we were in pain (although it really is), but that it *lacks a crucial ingredient*. Consequently, we (wrongly) intuit that *Pure Zombie Suffering* is not as bad as *Pure Suffering*. However, if we could *entirely and instinctively* believe illusionism, we would lose this intuition. In other words: we know that *our pain* (the one we really have, which is quasi-phenomenal) is bad, and at the same time we cannot help but believing (falsely) that our pain *is phenomenal*. We end up forming the belief that the (real) badness is grounded in its (illusory) phenomenality, which makes us believe falsely that

²⁰ To the best of my knowledge, the only views which have been developed and which entail the impossibility of consciousness also claim that the very concept of consciousness is incoherent (Dennett, 1988; Rey, 1995, 1997). That does not fit well with what I presupposed in this paper (i.e. the conceivability of consciousness). I know of no attempt to develop the view that consciousness is conceivable, but metaphysically impossible (which would require to deny modal rationalism: the thesis that what is conceivable is possible).

Phenomenal Value must be true, and without phenomenality, pain (for example) is not as bad as with phenomenality; but such belief would disappear if we became intuitive illusionists.

This strategy presents some advantages, but it also has drawbacks. It has notably the consequence that, at the times at which one fully accepts illusionism, the intuition supporting *Phenomenal Value* should disappear. So, it *discards* the intuition in favor of *Phenomenal Value* in a way that also presupposes that it should *disappear* under the right conditions (it mixes the “changing the intuition” and the “discarding the intuition” strategies). But that this intuition disappears when we are most convinced of the truth of illusionism is far from obvious – at least to me. One could then claim that there is a sense in which *no one* ever comes to fully embrace illusionism, because the illusion of phenomenality is too strong, but that would create some difficulty for illusionists – what does it even mean to be an illusionist if you cannot believe illusionism? Illusionists probably could solve this difficulty (for example, by distinguishing different ways in which one can *believe* a theory), but the resulting view might start to look *ad hoc* and/or untestable to some. In my mind, however, this last strategy probably remains the most attractive one for an illusionist who wants to deny *Phenomenal Value*.

7. Concluding remarks

There is an intuitive link between phenomenality and value. However, according to illusionists, phenomenality does not exist. This creates the *normative challenge* for illusionist views of consciousness. I think that illusionists should ultimately face up to this challenge. If we take illusionism to be true – or even merely to be one of the serious contenders amongst theories of consciousness – then we have to understand what it implies in the normative domain. If we embrace illusionism, should we revise our judgments concerning what is good and bad, what is better and worse? This, I think, is to a great extent an open (and fascinating) question;

but it is a question which, as I tried to show, is not easy to answer, as any stance towards the normative challenge encounters problems.

I have defended illusionism regarding consciousness elsewhere (Kammerer, 2016, 2019), and I still think that this is probably the true view of consciousness. One could then wonder which kind of stance towards the normative challenge I recommend. My goal here was not to defend a particular position towards the normative challenge (which would require more arguments), but rather to map the options. However, given what precedes, here is where my preference lies. My favorite option is the “moderate option”, as I find it hard to justify the rejection of *Phenomenal Value* in a satisfying way. Indeed, all the routes I examined to do so raise serious issues (even though I see a bit more hope in the one I called “*Skepticism about our value intuitions given our realist intuitions*”). I therefore think we should probably accept *some* revisionary normative consequences (admit that pain and pleasure are not as bad as thought, and weaken *Consciousness Contrast* and *Sentience Contrast*), while also providing views that justify and explain the value of valenced quasi-phenomenal states. The explanatory part will probably have to focus on the attitudes and desires that usually accompany or partly constitute valenced quasi-phenomenal states. As for the justificatory part, I think I am attracted by a second-person epistemology of the value of mental states (where empathy plays a crucial role), but I am not sure it can be worked out satisfactorily in the end. In fact, I suspect the justificatory part is going to correspond to the hardest part of the normative challenge for illusionism.

More generally, I am a bit suspicious of views that would manage to avoid revisionary consequences altogether. Revisionary normative consequences may often appear undesirable; but it would be very strange, after all, that a radical revision of our conception of what reality (in general) and the mind (in particular) *are* – and illusionism really is such a revision – should leave our conception of the *value* of the states of affairs (in general) and of states of mind (in particular) completely untouched.

Now, in conclusion, let us admit for one second that the best illusionist answer (or maybe all acceptable illusionist answers) to the normative challenge has deeply revisionary normative consequences (*nihilism* being one of the most striking examples of such deeply revisionary consequences). It would give us a *motivation* to reject illusionism, as we are usually resistant to normative revisions. However, will this also give us a *justification* to reject illusionism? In other words, does the normative challenge have epistemic upshots? This will depend on whether or not we have good *justifications* for our antecedent normative judgments. And here things may quickly become tricky, as illusionism itself might not only lead to *contradict* these antecedent normative judgments, but might also undercut their justification (for example if these antecedent normative judgments were supposed to be justified by introspection, the status of which, as a reliable justificatory process, is threatened by illusionism).

For the sake of the argument, let us now take our antecedent normative judgments to be independently justified. In this case, the normative challenge would potentially provide the basis of an argument against illusionism.²¹ How should the philosopher who has illusionist inclinations react?

One way to go would be to choose *weak illusionism* over *strong* illusionism (which I simply called “illusionism” in this paper). Weak illusionists claim that phenomenal consciousness exists, but does not have many of the problematic properties it is usually thought to have – it merely *seems* to have these properties. Weak illusionism does not deny the existence of phenomenality, which might seem to help in escaping revisionary normative consequences. However, I think that weak illusionism is at heart an unstable position: that it is either a merely *verbal* variation on strong illusionism, or that it is in fact incoherent, in the sense that it is implicitly committed to the existence of full-blown phenomenality (endowed with the kind of

²¹ For an argument against reductive *materialism* in that spirit, see (Cutter, 2017).

problematic properties that the weak illusionist explicitly denies).²² Another difficulty for weak illusionists in the context of the normative challenge stems from the fact that they believe we are the victims of illusions regarding phenomenality: phenomenal states seem to have essential properties they do not really have. But once you admit that our grasp of phenomenality is systematically mistaken in important respects, it becomes difficult to support the idea that the capacity of phenomenal states to ground some form of value is not itself amongst these illusory properties. In that sense, weak illusionists too might have to face revisionary normative consequences. One could also retreat to *reductive materialism*, which is a variety of phenomenal realism (but still seems closer in spirit to illusionism than, say, dualism or neutral monism). However, I take it that the most plausible versions of reductive materialism will be forms of weak illusionism, which means that they also face the difficulties just described.

Therefore, to the question “Can the normative challenge be used to argue against illusionism?” the answer is “Yes”. However, building such an argument will not be easy. Substantive suppositions are needed. Moreover, I doubt such an argument would comparatively give support to neighboring views, such as weak illusionism or reductive materialism: an argument against illusionism based on the normative challenge will probably be an argument against materialism in general (or at least against the most plausible forms of materialism).

Acknowledgments: I want to thank Romain Bourdoncle, David Chalmers, Keith Frankish, Thibaut Giraud, Uriah Kriegel, Andrew Lee, Tricia Magalotti, Johannes Martens, Luke Muehlhauser, Sonia Paz-Higgins, Andrea Peacocke, Henry Shevlin, Carl Shulman, Jonathan Simon, Cecily Whiteley for their helpful remarks and comments, as well as audiences at the Institut Jean Nicod, at the Université

²² In Keith Frankish’s terms (Frankish, 2012): there is no “dirt *qualia*” between “zero *qualia*” and full-blown *qualia*: there is no coherent conception of phenomenality which amounts to *more* than quasi-phenomenality but is still distinct from “classic” phenomenality. The weak illusionist is either a strong illusionist *or* a phenomenal realist in disguise.

Forthcoming in *Ergo*

Catholique de Louvain and at New York University. I also want to thank the two anonymous referees for their useful remarks.

References

- 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.
- Balog, K. (2012). Acquaintance and the Mind-Body problem. In C. Hill & S. Gozzano (Ed.), *New Perspectives on Type Identity: The Mental and the Physical*. Cambridge University Press.
- Block, N. (1995). On a confusion about a function of consciousness. *Brain and Behavioral Sciences*, 18, 227--247.
- Brandt, R. (1979). *A Theory of the Good and the Right*. Clarendon Press.
- Brentano, F. (1952). *Grundlegung und Aufbau der Ethik, Nach den Vorlesungen über Praktische Philosophie*. Meiner.
- Carruthers, P. (1999). Sympathy and subjectivity. *Australasian Journal of Philosophy*, 77(4), 465-482.
- Carruthers, P. (2004). Suffering without subjectivity. *Philosophical Studies*, 121(2), 99-125.
- Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. (2007). Phenomenal concepts and the explanatory gap. In T. Alter & S. Walter (Ed.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford University Press.
- Chalmers, D. (2018). The Meta-Problem of Consciousness. *Journal of Consciousness Studies*, 25(9-10), 6-61.

Forthcoming in *Ergo*

Crisp, R. (2006). *Reasons and the Good*. Oxford University Press.

Cutter, B. (2017). The Metaphysical Implications of the Moral Significance of Consciousness. *Philosophical Perspectives*, 31(1), 103-130.

Demircioğlu, E. (2013). Physicalism and Phenomenal Concepts. *Philosophical Studies*, 165(1), 257-277.

Dennett, D. (1988). Quining Qualia. In A. Marcel & E. Bisiach (Ed.), *Consciousness in Modern Science*. Oxford University Press.

Dennett, D. (1991). *Consciousness Explained*. Penguin.

Dennett, D. (2016). Illusionism as the Obvious Default Theory of Consciousness. *Journal of Consciousness Studies*, 23(11-12), 65-72.

Dennett, D. (2017). *From Bacteria to Bach and Back*. Norton & Company.

Diaz-León, E. (2014). Do a Posteriori Physicalists Get Our Phenomenal Concepts Wrong? *Ratio*, 27(1), 1-16.

Elpidorou, A. (2013). Having it both ways: consciousness, unique not otherworldly. *Philosophia*, 41(4), 1181-1203.

Elpidorou, A. (2016). A posteriori physicalism and introspection. *Pacific Philosophical Quarterly*, 97(4), 474-500.

Feldman, F. (2002). The Good Life: A Defense of Attitudinal Hedonism. *Philosophy and Phenomenological Research*, 65, 604-627.

Feyerabend, P. (1963). Mental Events and the Brain. *Journal of Philosophy*, 40, 296-296.

Fine, K. (2012). Guide to Ground. In F. Correia & B. Schnieder (Ed.), *Metaphysical Grounding*. Cambridge University Press.

Frankena, W. (1973). *Ethics* (2nd Edition). Prentice Hall.

Frankish, K. (2012). Quining diet qualia. *Consciousness and Cognition*, 21(2), 667-676.

- Frankish, K. (2016). Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies*, 23(11-12), 11-39.
- Glannon, W. (2016). The value and disvalue of consciousness. *Cambridge Quarterly of Healthcare Ethics*, 25(4), 600-612.
- Goff, P. (2011). A Posteriori Physicalists Get Our Phenomenal Concepts Wrong. *Australasian Journal of Philosophy*, 89(2), 191-209.
- Graziano, M. (2013). *Consciousness and the Social Brain*. Oxford University Press.
- Heathwood, C. (2006). Desire satisfactionism and hedonism. *Philosophical Studies*, 128, 539-563.
- Irvine, L., & Sprevak, M. (forthcoming). Eliminativism about consciousness. In U. Kriegel (Ed.), *Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32(April), 127-136.
- Joyce, R. (2006). *The Evolution of Morality*. MIT Press.
- Kahane, G. (2009). Pain, Dislike and Experience. *Utilitas*, 21(3), 327-336.
- Kammerer, F. (2016). The hardest aspect of the illusion problem - and how to solve it. *Journal of Consciousness Studies*, 23(11-12), 123-139.
- Kammerer, F. (2019). The illusion of conscious experience. *Synthese*.
<https://doi.org/10.1007/s11229-018-02071-y>
- Kriegel, U. (forthcoming). The Value of Consciousness: A Propaedeutic. *Analysis*.
- Kriegel, U. (2009). *Subjective Consciousness : A Self-Representational Theory*. Oxford University Press.
- Kriegel, U. (2018). *Brentano's Philosophical System*. Oxford: Oxford University Press.
- Lee, A. (2018). Is consciousness intrinsically valuable? *Philosophical Studies*, 175(1), 1-17.
- Lee, G. (2014). Materialism and the Epistemic Significance of Consciousness. In U. Kriegel (Ed.), *Current Controversies in Philosophy of Mind* (p. 222-245). Routledge.

Forthcoming in *Ergo*

Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. Oxford University Press.

Levine, J. (2007). Phenomenal Concepts and the Materialist Constraint. In T. Alter & S. Walter (Ed.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford University Press.

Levy, N. (2014). The Value of Consciousness. *Journal of Consciousness Studies*, 21(1-2), 127-138.

Levy, N., & Shepherd, J. (forthcoming). Consciousness and morality. In U. Kriegel (Ed.), *The Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.

Loar, B. (1997). Phenomenal States (Revised Version). In N. Block, O. Flanagan, & G. Güzeldere (Ed.), *The Nature of Consciousness* (p. 597-616). MIT Press.

Mackie, J. (1977). *Ethics: Inventing Right and Wrong*. Penguin.

Moore, A. (2013). Hedonism. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition).

Muehlhauser, L. (2017). Report on Consciousness and Moral Patienthood. *Open Philanthropy*. <https://www.openphilanthropy.org/2017-report-consciousness-and-moral-patienthood>

Nagel, T. (1979). *Mortal Questions*. Cambridge University Press.

Nida-Rümelin, M. (2007). Grasping Phenomenal Properties. In T. Alter & S. Walter (Ed.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford University Press.

Papineau, D. (2002). *Thinking about Consciousness*. Oxford University Press.

Pereboom, D. (2011). *Consciousness and the Prospects of Physicalism*. Oxford University Press.

Rey, G. (1995). Towards a Projectivist Account of Conscious Experience. In T. Metzinger (Ed.), *Conscious Experience*. Ferdinand Schöningh.

Forthcoming in *Ergo*

Rey, G. (1997). A Question about Consciousness. In N. Block, O. Flanagan, & G. Güzeldere (Ed.), *The Nature of Consciousness* (p. 961-982). MIT Press.

Rorty, R. (1965). Mind-Body Identity, Privacy and Categories. *Review of Metaphysics*, 19, 24-54.

Schroer, R. (2010). Where's the Beef? Phenomenal Concepts as Both Demonstrative and Substantial. *Australasian Journal of Philosophy*, 88(3), 505-522.

Seager, W. (2001). Consciousness, Value and Functionalism. *Psyche*, 7, 1-19.

Sebo, J. (2017). Agency and Moral Status. *Journal of Moral Philosophy*, 14(1), 1-22.

Shea, N. (2014). Using phenomenal concepts to explain away the intuition of contingency. *Philosophical Psychology*, 27(4), 553-570.

Shepherd, J. (2016). Moral Conflict in the Minimally Conscious State. In W. Sinnott-Armstrong (Ed.), *Finding Consciousness: The Neuroscience, Ethics, and Law of Severe Brain Damage*. Oxford University Press.

Sidgwick, H. (1907). *The Method of Ethics*. Thoemmes Press.

Siewert, C. (1998). *The Significance of Consciousness*. Princeton University Press.

Siewert, C. (2014). Speaking Up For Consciousness. In U. Kriegel (Ed.), *Current Controversies in Philosophy of Mind*. Routledge.

Singer, P. (2011). *Practical Ethics, 3rd Edition*. Cambridge University Press.

Sinhababu, N. (manuscript). The epistemic argument for hedonism.

<https://philpapers.org/archive/SINTEA-3.pdf>

Sprigge, T. (1988). *The Rational Foundations of Ethics*. Routledge.

Strawson, G. (2018). The Consciousness Deniers. *New York Review of Books*, March 13.

<https://www.nybooks.com/daily/2018/03/13/the-consciousness-deniers/>

Street, S. (2006). A Darwinian dilemma for realist theories of value. *Journal of Philosophy*, 91, 221-235.

Forthcoming in *Ergo*

Streumer, B. (2013). Can we believe the error theory? *Journal of Philosophy*, 110, 194-212.

Sturgeon, S. (1994). The Epistemic View of Subjectivity. *The Journal of Philosophy*, 91(5), 221-235.

Tännsjö, T. (1998). *Hedonistic Utilitarianism*. Edinburgh University Press.

van der Deijl, W. (2019). Is pleasure all that is good about experience? *Philosophical Studies*, 171(7), 1769-1787.

Zimmerman, M. (2015). Intrinsic vs. extrinsic value. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition).