Desire-Fulfilment and Consciousness

Andreas Mogensen (Global Priorities Institute, University of Oxford)

Global Priorities Institute | October 2024

GPI Working Paper No . 24-2024

Please cite this working paper as: Mogensen, A. Desire-Fulfilment and Consciousness. *Global Priorities Institute Working Paper Series, No. 24-2024.* Available at https://globalprioritiesinstitute.org/desire-fulfilment-and-consciousness-andreas-mogensen

GLOBAL PRIORITIES INSTITUTE

UNIVERSITY OF OXFORD

Desire-Fulfilment and Consciousness

Andreas L. Mogensen

Abstract

I show that there are good reasons to think that some individuals without any capacity for consciousness should be counted as welfare subjects, assuming that desire-fulfilment is a welfare good and that any individuals who can accrue welfare goods are welfare subjects. While other philosophers have argued for similar conclusions, I show that they have done so by relying on a simplistic understanding of the desire-fulfilment theory. My argument is intended to be sensitive to the complexities and nuances of contemporary developments of the theory, while avoiding highly counter-intuitive implications of previous arguments for the same conclusion.

1 Introduction

An individual has moral standing just in case they matter morally in their own right and for their own sake (Kamm 2007: 227–230). You and I have moral standing. Rocks do not. Why not?

According to one influential view, it's because rocks aren't sentient. 'Sentience' can be understood in both a broad and narrow sense (Browning and Birch 2022). In the broad sense, it's synonymous with the capacity for phenomenal consciousness. In the narrow sense, which I'll rely on here, sentience requires not only the capacity for phenomenal consciousness, but for phenomenal states that feel good or bad, exemplified by experiences of pleasure and pain.

According to Singer (1993), sentience (so understood) is a necessary condition for moral standing, because it's a necessary condition for being the kind of individual whose life can go better or worse. Stated otherwise, sentience is necessary for being a *welfare subject*. Singer's overall view might be summarized as the claim that welfare subjectivity is necessary and sufficient for moral standing and sentience is necessary and sufficient for welfare subjectivity.

In this paper I'll grant the first biconditional stated above for the sake of argument and focus

on the following challenge to the second. The assumption that sentience is a prerequisite for well-being appears to be in tension with a bunch of theories of welfare, as a number of philosophers have noted recently (Kagan 2019: 32–34; van der Deijl 2020; Bradford 2022).[1] Notably, that includes the *desire-fulfilment theory*, which Singer (1993) endorsed,[2] since it's far from obvious that consciousness is a necessary condition for having satisfied desires.

Should we think that some individuals without any capacity for phenomenal consciousness can be benefited and harmed because they have desires that may or may not be fulfilled? How we answer this question may have implications that reach beyond the desire-fulfilment theory itself. Objective list theories of welfare may recognize desire-fulfilment as one welfare good among others (e.g., Arneson 1999). Hybrid theories may treat fulfilment of desires for particular objectively valuable objects as constitutive of welfare (Parfit 1984: 501–502; Kraut 1994), raising the question of whether individuals without any capacity for phenomenal consciousness can instantiate these objective-subjective hybrids.

In this paper, I'll show that there are indeed good reasons to think that some individuals should be counted as welfare subjects although they lack the capacity for consciousness, provided we assume that desire-fulfilment is a welfare good and that any individuals who can accrue welfare goods are welfare subjects. While other philosophers have argued for similar conclusions, I'll show that they have done so by relying on a simplistic understanding of the desire-fulfilment theory. My argument is intended to be sensitive to the complexities and nuances of the theory and to avoid certain highly counter-intuitive implications that arise from reliance on a simplistic understanding of its commitments.

I begin in section 2 by explaining why some philosophers have thought that the desire-fulfilment theory of welfare requires us to recognize the possibility of welfare subjects without any capacity for phenomenal consciousness and why I think those arguments are unconvincing. In section 3, I consider a plausible modification of the desire-fulfilment theory put forward by Sumner (1996) and Heathwood (2019) and argue that the best way to understand this proposal treats it as recognizing positive affect as necessary for the welfare good of desire-fulfilment. Whereas Lin

[1] See also Berger et al. (ms) and Goldstein and Kirk-Giannini (ms).

[2] Singer now endorses a hedonistic theory of welfare instead (de Lazari-Radek and Singer 2014).

(2020a,b) argues that the kind of modified desire-fulfilment theory discussed in section 3 makes phenomenal consciousness a necessary condition for being a welfare subject, I take the contrary view. Although many find it intuitive to suppose that affective states have to be consciously experienced, in section 4, I show that there are good grounds for recognizing the possibility of affective states occurring without phenomenal consciousness. In section 5, I argue that there are also good grounds for according moral standing to individuals without any capacity for phenomenal consciousness who exhibit emotions and sentiments. I conclude that we have good reason to think that the putative welfare good of desire-fulfilment can be instantiated by individuals without any capacity for consciousness, although the conditions are far more restrictive than the arguments I'll discuss in section 2 would suggest. In section 6, I explore some of the implications of the view developed here for our thinking about invertebrates and AI systems as welfare subjects.

2 Desire-Fulfilment and Consciousness

The desire-fulfilment theory is one of three major theories of well-being (Parfit 1984: 493–502). Very roughly, it says that welfare consists in desire-fulfilment.[3] This is only a very rough statement of the theory, which virtually no one accepts without some additional qualification. For example, we may want to introduce a constraint that only the satisfaction of intrinsic desires is good in itself for the desirer (Brandt 1979: 111; Parfit 1984: 117).[4] We might also impose the constraint that it's only those desires whose objects are states of affairs that have the desirer as an essential constituent whose satisfaction is good in itself for the desirer (Overvold 1982; compare Parfit 1984: 494) or only those desires that would be held under idealized epistemic conditions (Rawls 1971: 416–417; Brandt 1979: 268).

Intuitively, desires themselves need not be conscious. At any given point in time, there are a great many things that I want and most of these wants make no difference to the character of my experience. For example, it's only when some specific trigger calls them to mind that I become

[3] Arguably, this is only a very rough statement of what Bradley (2014) calls the 'combo view,' to be contrasted with the 'object view,' on which welfare consists in those things that satisfy our desires, rather than the composite of the object of desire and our desire for it. However, I am convinced by Lin (2022) that this is a distinction without a difference.

[4] An intrinsic desire is, roughly, one whose object is not desired merely for the sake of some further end.

consciously aware of my desire to take a holiday in the Seychelles or to finish reading Iain M. Banks' *Culture* series. The natural view is that I have wanted those things all along, although these desires have figured only intermittently as constituents of my conscious mental life throughout the time I have had them.

It's a controversial question in the philosophy of mind whether propositional attitudes and other states with intentional content can also occur without being even potentially conscious or otherwise appropriately related to conscious states of the same subject (Searle 1990, 1992; Horgan and Tienson 2002; Loar 2003; Smithies 2012; Bourget and Mendelovici 2019). Still, I take it to be the mainstream view that they can, and I'll rely on that assumption throughout this paper. According to one influential view, desires are simply states of the agent that combine appropriately with the agent's belief-states to produce behaviour conducive to realizing the desired state of affairs, doing so in ways that make no explicit reference to any essential role for phenomenal consciousness (Stalnaker 1984: 1–26; Smith 1987; Schroeder 2004: 11–15).

If phenomenal consciousness is not a necessary condition for having desires, it might then seem to follow from the desire-fulfilment theory of welfare that individuals with no capacity for phenomenal consciousness can be welfare subjects, since they can have desires that may or may not be fulfilled. In fact, this seems to follow just so long as desire-fulfilment is *a* welfare good, even if it is only one among many others, as opposed to the one and only welfare good, as it is according to the desire-fulfilment theory. Kagan (2019: 33) and Bradford (2022) both set out this form of argument and recommend that we follow where it leads, acknowledging the possibility of unconscious welfare subjects.

An alternative response is to treat the argument as a *reductio* (compare van der Deijl 2020). It might especially concern us that following the argument where it leads may require us to recognize corporations as welfare subjects with moral standing (compare Goldstein and Kirk-Giannini ms). Corporations are often described as wanting to achieve various aims, and there is evidence that people treat ascriptions of propositional attitudes to corporations as literally true (Arico et al. 2014). Moreover, there are good philosophical reasons to take those ascriptions literally and to see corporations as capable of being guided by beliefs and desires that belong to the corporate entity itself, rather than its individual members (Clark 1994; Copp 2006; Pettit 2007; List and Pettit 2011; Hess 2014). Nonetheless, it may strike us as incredible to suppose that corporations are welfare

subjects who matter morally in their own right and for their own sake (Manning 1984; Wringe 2014; Hess 2018; List 2021; but see Silver 2019). We do not feel sympathy for Samsung.

Rather than following the argument where it leads or treating it as a *reductio*, we might instead deny that Kagan and Bradford reason correctly about the implications of the desire-fulfilment theory. For example, we could question the assumption that any individuals capable of possessing the welfare goods identified by a given theory of welfare should be counted as welfare subjects according to that theory (Lee forthcoming). Perhaps desire-fulfilment is what makes someone's life go better or worse, given that they are the kind of individual whose life can go better or worse, but what makes someone the kind of individual whose life can go better or worse is something else, like sentience.

I want to grant the assumption that any individuals capable of possessing welfare goods are welfare subjects and focus on a different challenge to the argument set out by Kagan and Bradford. Recall that almost no one accepts the bare theory that welfare consists in the fulfilment of just any kind of desire. A view of that kind is highly implausible and subject to a wide range of apparent counter-examples (Parfit 1984: 494–499). That's why various restrictions on the kinds of desires whose fulfilment contributes to welfare are typically thought to be necessary, such as the restriction to intrinsic desires and/or to desires for states of affairs that involve the desirer as an essential constituent.

It follows that having desires doesn't entail that you're a welfare subject on the desire-fulfilment theory, even granting that any individuals capable of possessing the welfare goods identified by a given theory of welfare are welfare subjects according to that theory. In addition, your desires need to satisfy the relevant supplementary restrictions required for the theory to be credible. A cogent argument demonstrating that the desire-fulfilment theory requires us to acknowledge the possibility of welfare subjects without any capacity for phenomenal consciousness needs to show that that capacity is also unnecessary for satisfying these supplementary restrictions.

We could imagine a form of the desire-fulfilment theory that simply introduces sentience or phenomenal consciousness as a self-standing supplementary restriction. This obviously has an *ad hoc* quality, although it's not obviously more *ad hoc* than the other restrictions that have been added to the theory in response to various purported counter-examples. Nonetheless, I set this view aside. Instead, I focus on a proposal for adding bells and whistles to the desire-fulfilment theory

put forward by Sumner (1996) and Heathwood (2019) independently of recent debates about the relationship between consciousness, welfare subjectivity, and moral standing.[5] In the next section, I set out their proposal and explain my own preferred interpretation of the dichotomy on which it relies. I then return to the issue of phenomenal consciousness in section 4.

3 Behavioural and Attitudinal Desire

Sumner and Heathwood argue that in order to develop a plausible desire-fulfilment theory of welfare, we need to distinguish two different senses of 'desire'.

In one sense, a person can be understood as desiring something just in case they are motivated to bring it about, with no restriction on the possible range of reasons, if any, that they might have for trying to achieve a given outcome. Call this *behavioural desire*.[6] This is arguably the standard way desire has been understood in analytic philosophy (Schroeder 2004: 11–27). However, it's not the only conception of desire available to us.

Note, for example, that when desire is understood as behavioural desire, a person cannot voluntarily choose to do anything they don't want to. However, people sometimes describe their actions in this way. For example, former US president George H. W. Bush describes agreeing to some demands made by the Democratic leadership as follows: "I didn't want to, but felt I had to, to get this deficit down" (Bush 1990: 1314). Clearly, Bush had in mind a different sense of 'want' than is captured by the merely behavioural interpretation.

What is this other sense? Sumner (1996: 121) suggests that there is a different sense of 'want' on which wanting something "requires finding the prospect of it pleasing or agreeable, or welcoming the opportunity to do it, or looking forward to it with gusto or enthusiasm."[7] I take it Bush didn't feel that way about compromising with his political opponents. Let's call the sort of pro-attitude

[5] For prior discussion of the same proposal in the context of debates about the relationship between consciousness, welfare subjectivity, and moral standing, see Lin (2020b) and Goldstein and Kirk-Giannini (ms).

[6] As opposed to following either author's nomenclature strictly, the particular terms I use here are something of a mash-up of the different terms Sumner and Heathwood use.

[7] Compare Heathwood (2019: 674): "the person finds the occurence of the event attractive or appealing, is enthusiastic about it (at least to some extent), and tends to view it with pleasure or gusto".

Bush was missing *attitudinal desire*.

Sumner and Heathwood argue that it is only attitudinal desire whose fulfilment is a plausible candidate for making a person better off. Heathwood makes an especially strong argument, showing that this hypothesis allows us to respond to a variety of cases that otherwise seem to pose serious problems for desire-fulfilment theories. For example, it avoids the result that people we would naturally describe as sacrificing their own self-interest do no such thing, since they do what they most strongly desire to do. Suppose you decline to eat the last slice of pizza, although it looks delicious. You don't want to be the person to take the last slice. You get what you want. Still, the intuitive thing to say is that you sacrificed your our own self-interest for the sake of etiquette. This can be explained straightforwardly by appeal to the claim that only the satisfaction of attitudinal desire is a welfare good, given the plausible assumption that standing on principles of etiquette isn't the sort of thing that inspires gusto or enthusiasm, unlike delicious pizza (Heathwood 2019: 677–678).

But what exactly is attitudinal desire? We are told that attitudinal desire may involve finding the desired object pleasing or agreeable and that it may involve viewing it with pleasure or gusto. The implication is that there are different ways this sort of attitude can manifest. What is it that manifests?

A very plausible answer, I take it, is that attitudinal desire involves a (disposition to exhibit some) positive affective state directed toward the object of desire (compare Vadas 1984). What do I mean by an affective state? Affective states are a class of psychological states of which emotions are the paradigm instance. Other commonly recognized affective states include moods and valent bodily states like itches and pains.[8]

What do these different states have in common that mark them out as a distinctive psychological kind? I think that is an open empirical question. James A. Russell's influential circumplex model characterizes affective states in terms of their level of *arousal* and their *valence* (Russell 1980, 2003). The valence of an affective state here refers to its position along a positive-negative axis, with emotions like joy or contentment falling toward the positive end and anger or fear toward the negative. However, there is no agreed-upon theory of what the valence of affective states consists

[8]Some classify these as just emotions of a particular kind: Prinz (2004: 182–188) argues that moods are just a special class of emotions; Craig (2003) counts itches and pains as 'homeostatic emotions'.

in (see Prinz 2004: 160–178; Carruthers 2017; Barlassina and Hayward 2019; Martínez and Barlassina forthcoming). A wide range of authors who otherwise accept diverse theories of emotion converge on the idea that affective states involve interoceptive registrations of bodily states and physiological changes (James 1884; Lazarus 1991; Damasio 1994, 1999; Prinz 2004; Barrett 2017). However, others deny a central role for registrations of bodily states and changes in explaining affect (Nussbaum 2001; Rolls 2013).

I won't take a stand on these controversial issues. To simplify discussion, I'll largely focus on the role of positive emotions in grounding attitudinal desire, setting aside other affective states.

One defining feature of emotions – both positive and negative – is that they are *intentional states* (Pitcher 1965; Deonna and Teroni 2012: 3–6). Emotions are directed toward intentional objects, which may or may not exist, as when a child is afraid of the monster under the bed. Intentionality may not be a property shared by all affective states. For example, Nussbaum (2001: 132–5) argues that moods differ from emotions precisely in that moods do not have intentional objects. If so, moods cannot be directed toward objects of desire, and hence cannot ground attitudinal desires. Pains and itches have also struck philosophers including Rorty (1970) and Searle (1979) as lacking intentional content, although that view is now widely rejected (see Tye 1995; Bain 2003; Klein 2007; Cutter and Tye 2011).

I take it that the hypothesis that attitudinal desire involves a (disposition to exhibit some) positive affective state directed toward the object of desire is highly plausible on its face. It also explains why attitudinal desire can manifest in the wide variety of different ways noted by Sumner and Heathwood. As I've observed, affective states come in different varieties. Positive emotions themselves vary considerably in their degree of valence, level of arousal, and characteristic mode of expression in thought, experience, and action, as can be seen by comparing positive emotions such as joy, contentment, and awe.

The proposal also lends itself naturally to the development of a theory of *ill-being*, i.e., a theory of what makes someone's life go badly for them (Kagan 2014). Some lives are not merely void of good things. They are filled up with things that are bad in themselves. What account can the desire-fulfilment theory give of these prudential bads? A natural idea is to posit an attitude – *aversion* – that is the opposite of desire. When the state of affairs that is the object of this negative attitude obtains, that is bad in itself for the individual, just as it is good in itself when a desired

state of affairs obtains (Kagan 2014; Sumner 2020; Pallies 2022; Heathwood 2022).

The current proposal suggests a natural account of the nature of aversion and makes it clear why positing an attitude of this kind isn't *ad hoc*. As noted, one of the hallmarks of affective states is that they are valenced and can be arranged along a positive-negative axis. Granting that attitudinal desire involves a positive affective state, we can then say that aversion involves a corresponding negative affective state.[9] Admittedly, this only gets us so far. As I've already noted, it is an open question how exactly we should account for the valence of affective states. What is it that makes some emotions positive and others negative? The current suggestion is that we can pass the buck to theorists of emotion, explaining the opposing valence of desire and aversion in whatever terms best explain the nature of emotional valence.

4 Attitudinal Desire and Consciousness

We started off with the following idea. Suppose that desire-fulfilment is a welfare good. Desires arguably needn't be conscious or even potentially conscious. It seems to follow that the capacity for phenomenal consciousness is unnecessary for being a welfare subject.

Based on the discussion so far, we have reason to suspect that this argument trades on a fallacy of ambiguity. The claim that desire has no essential connection to consciousness seems plausible if by 'desire' we mean 'behavioural desire.' However, it isn't plausible that the satisfaction of merely behavioural desires is a welfare good. On the other hand, the claim that desires need not be conscious or potentially conscious is much less plausible if 'desire' is understood to mean 'attitudinal desire'.

In the recent philosophical literature, Lin (2020a,b) suggests that attitudinal desire requires phenomenal consciousness. He notes that although Sumner and Heathwood do not explicitly say so, "it is plausible that [attitudinal desires] are partly constituted by a certain kind of phenomenology: to desire something in the [attitudinal] sense is, among other things, to feel a certain way about

[9] Compare Pallies (2022: 618): "attraction involves a certain sort of directed anticipatory pleasure; aversion involves a certain sort of directed anticipatory displeasure." If we understand affective valence as a matter of felt (un)pleasantness, this plausibly amounts to more or less the same idea. However, for reasons that will become clear, I think we should reject a hedonic theory of valence.

it." (Lin 2020a: 520)

My own proposal is in some ways similar. I claim that attitudinal desire should be understood as involving a (disposition to exhibit some) positive affective state, such as an emotion. To have a positive emotion is naturally described as to feel a certain way about something: namely, to feel positively about it. Many also find it intuitive to suppose that emotions have some kind of necessary connection with phenomenal consciousness. We might think the way an emotion colours conscious experience is essential to it – or even exhausts its nature. William James (1884) famously takes this view. Even Freud (1915 / 1963: 126) writes that "for emotions, feelings, and affects to be unconscious would be quite out of the question."

Nonetheless, the view that emotions must be consciously experienced has many detractors among contemporary philosophers, psychologists, and neuroscientists (Damasio 1999; Prinz 2004: 201–205; Berridge and Winkielman 2003). Scarantino (2010: 738) describes the view that emotions must have experiential qualities as "at this point a minority view on the relation between emotion and feeling." Therefore, if the defining feature of attitudinal desire is the involvement of positive affect, the relationship between attitudinal desire and phenomenal consciousness is likely to be a good deal more contested than Lin suggests (compare Berger et al. ms).

Why think that affective states can occur outside phenomenal consciousness? The strongest case for the existence of unconscious emotion would take the form of a successful psychological theory identifying emotion with a psychological kind that cross-cuts the distinction between conscious and unconscious mental processing (Adolphs and Anderson 2018: 49–51). We have theories that make that kind of identification, although their success is debatable.

For example, Prinz (2004) outlines a theory of emotions as *embodied appraisals* and notes that the functions associated with emotional processing on this theory are not intrinsically bound to consciousness. On Prinz's view, emotions register physiological responses, such as changes in heart rate and blood flow. In doing so, their ultimate function is to indirectly track core relational themes, such as the presence of danger or the occurrence of loss. Since neither the registration of bodily changes nor the tracking of core relational themes appears to require phenomenal consciousness, phenomenal consciousness is not a necessary condition for emotion on Prinz's account (see Prinz 2004: 198–220).

Although I think Prinz's account has much to recommend it, resting the argument for the possi-

bility of unconscious emotion on this kind of fully-fledged theory of emotion is unlikely to persuade very many. Fortunately, an argument can also be made without appeal to any controversial theory of emotion, relying instead on intuitions about cases.

A number of different purported examples of unconscious emotion have been put forward.[10] For example, people undergoing an emotional response sometimes appear to be completely distracted from their own emotional state, their attention fully occupied by something else (see Pitcher 1965: 338; Goldie 2002: 62; Prinz 2004: 201–202). The driver of an out-of-control vehicle may be completely absorbed in ensuring they do not crash while apparently gripped by fear. They might not realize how terrified they were throughout the ordeal until the car comes to a stop and they notice how far their fingernails have dug into the steering wheel. If we adopt the (admittedly controversial) assumption that attention gates the contents of consciousness (Baars 1988; Dehaene et al. 2006; Prinz 2012), this sort of case might be thought of as involving unconscious emotion.

Results from a series of studies by Kent Berridge and Piotr Winkielman also provide evidence for the occurrence of unconscious affective states (Berridge and Winkielman 2003; Winkielman and Berridge 2004; Winkielman et al. 2005).[11] Unlike in the kind of example noted in the previous paragraph, these are unconscious affective states that people aren't able to introspect even when asked to attend to how they're feeling.

Berridge and Winkielman's studies involve subliminal presentation of a happy, neutral, or angry face. Prior studies had shown that subliminal images can evoke emotions that are registered by subjects' self-reports (Öhman and Soares 1994). In Berridge and Winkielman's studies, the prime did not influence self-reported affect.[12] However, it did influence behaviour. They recorded

---

[10] I set aside purported cases of unconscious emotion that can easily be accommodated by distinguishing between occurrent emotions and emotional dispositions. For example, Solomon (1973: 23) writes that "it is clear that one can have an emotion without feeling anything. One can be angry without feeling angry: one can be angry for three days or five years and not feel anything identifiable as a feeling of anger continuously throughout that prolonged period." As Solomon concedes, without further argument, these observations can be accommodated by positing that being angry is here simply a disposition to experience feelings of anger. The examples I consider all involve occurrent affective states and so can't be dismissed in this way.

[11] See Winkielman and Gogolushko (2018) for a recent replication.

[12] A common worry about Berridge and Winkielman's results is that their subjects might simply have undergone changes

subjects' reactions to a lemon-lime flavoured drink. The valence of the prime influenced thirsty subjects' reactions to the drink, as revealed by how much of the beverage they poured and consumed. The authors take this to provide "the first demonstration in a non-clinical, human sample that affective reactions can be subliminally triggered and can change behavior yet still remain inaccessible to introspection." (Winkielman et al., 2005: 132)

Not everyone will come away convinced from this brief survey of reasons to recognize the possibility of unconscious affective states (see Hatzimoysis 2007, Whiting 2018). Sadly, a deeper and more thorough-going discussion would take us too far afield. Nonetheless, I think the considerations I've reviewed make a reasonable case for recognizing the possibility of affective states occurring outside of consciousness. *Ipso facto*, they make a reasonable case for supposing that there can be attitudinal desires in the absence of consciousness. At the very least, they suggest that this is a hypothesis worth taking seriously. The issue I want to focus on now concerns the normative implications of doing so.

5 Unconscious Affect and Moral Standing

In section 3, I argued that the welfare good of desire-fulfilment requires the participation of positive affective states. Whereas we might have thought that this implied a necessary role for phenomenal consciousness, I argued in the previous section that we have good reason to expect that our best psychological theories might fail to treat affective states as intrinsically bound to consciousness.

Even if a policy of reserving the term 'emotion' for consciously experienced mental states fails to carve psychological reality at the joints, it doesn't follow that it also fails to carve ethical reality at the joints. Why not simply insist on drawing yet another distinction here, this time within the class of attitudinal desires and affects? Why not say that it is only attitudinal desires that involve *conscious* affective states whose satisfaction contributes to welfare?

Consider again the kind of affective reactions evoked by Berridge and Winkielman in their subjects as a result of the subliminal presentation of an angry or happy face. Doesn't it seem hard to believe that those unconscious responses made any difference at all to people's welfare?

---

in consciously experienced affect of a kind too subtle or short-lived to be noticed (Whiting 2018; Birch 2024: 99). See Prinz (2004: 202–205) and Winkielman et al. (2005: 132) for replies to this concern.

Perhaps so. Nonetheless, these affective reactions are arguably so weak in themselves that any contribution they make to individual welfare might be similarly easy to miss. A better test-case would involve strong unconscious emotions similar to intense fear or anger. Unsurprisingly, there are no documented experiments in which emotions like that are subliminally induced in human subjects (see Adolphs and Anderson 2018: 304–306). Fortunately, we're philosophers. We can appeal to thought-experiments.

Consider the following hypothetical case due to Kagan (2019). Imagine that in the distant future, we discover another planet, on which there exists a civilization of robots. Relying on our best scientific theory of phenomenal consciousness, we learn that these robots do not have any capacity for conscious experience. Nonetheless, we are asked to imagine that they engage in a variety of sophisticated cognitions and behaviours. They make art. They make friends. They even reproduce mechanically and form families.

Kagan asks us to imagine that you're a scientist hoping to learn more about the inner workings of these strange creatures:

> So you capture a small one – very much against its protests – and you are about to cut it open to examine its insides, when another robot, its mother, comes racing up to you, desperately pleading with you to leave it alone. She begs you not to kill it, mixing angry assertions that you have no right to treat her child as though it were a mere thing, with emotional pleas to let it go before you harm it any further. (Kagan 2019: 28)

Kagan thinks that it's obvious that it would be wrong to rip up the small robot you have captured. "It simply doesn't matter to me that the child and its mother are 'mere' robots, lacking in sentience." (Kagan 2019: 28)

I share the intuition that it would be wrong to destroy the child. However, I don't agree with the lesson Kagan draws from this. According to Kagan, being an *agent* suffices for moral standing. To be an agent, in Kagan's sense, is to have self-regarding preferences that guide behaviour. Kagan understands these preferences as amounting to behavioural desires: "dispositions to act in ways that tend – according to one's beliefs – to bring about specified states of affairs." (Kagan 2019: 20) He denies that the mental states constitutive of agency require phenomenal consciousness and suggests that even if they do, there must nonetheless exist analogous states with similar functional roles that we should think of as sufficient for moral standing.

I find it implausible that our reaction to Kagan's thought-experiment is driven by the presence of agency, so understood. Agency, as Kagan understands it, is something that the robots in his fable seem to share with corporations.[13] Breaking up a corporation does not evoke the kind of reaction we have to the prospect of breaking up the small robot in Kagan's thought experiment. What does the work in this thought experiment is surely the attribution to that robot of a kind of mental state that corporations saliently lack (*pace* Gilbert 2002): namely, emotion. What grabs us is the fact that the child *protests* and that the mother pleads *desperately* and makes *angry* assertions.

If we are willing to take our intuitions about this case at face value, the conclusion it most naturally suggests is that *agency infused with emotion* suffices for moral standing in the absence of a capacity for phenomenal consciousness. This is exactly what we should expect if desires backed by unconscious affective states can participate in the welfare good of desire-fulfilment, given the other assumptions we have taken onboard.

I concede that we should be cautious in relying on our intuitions about this case. It's natural to worry about our ability to form a coherent picture of what is supposed to be going on insofar as we are asked to imagine both that the robots in the thought-experiment are unconscious and that they undergo strong emotions, since we intuitively associate emotion with conscious feelings. I think that's a perfectly reasonable concern. For this reason, I think it would be a mistake to put a great deal of weight on our intuitive responses to Kagan's imagined scenario. But that is not the same as giving them no weight at all.

Kagan's thought experiment may well be the best we can do in constructing a critical test of the intuitive plausibility of the claim that desires backed by affective states can suffice for moral standing even in the absence of any capacity for phenomenal consciousness. Reflecting on this case suggests no reason to complicate our theory by distinguishing among the welfare significance of attitudinal desires in the way proposed at the start of this section and gives us some positive reason (beyond mere parsimony) to insist that the more complex theory is unlikely to be correct.

A different kind of thought experiment may be thought to challenge the necessity of affective states for moral standing. Chalmers (2022: 339-344) asks us to imagine *philosophical Vulcans*, or *p-Vulcans*, for short. P-Vulcans are conscious creatures that are very much like human beings,

---

[13] Obviously, they are also one another's phenomenal duplicates: they trivially share all the same phenomenal properties.

except that they are not sentient (in the narrow sense) and so experience no happiness, suffering, pleasure, or pain. They nonetheless have serious intellectual and moral goals, although they take no pleasure in achieving those goals and experience no felt unpleasantness at their frustration.

Chalmers appeals to p-Vulcans to challenge Singer's claim that moral standing requires sentience. If standing requires sentience, he notes, this would seem to entail the permissibility of killing arbitrarily many p-Vulcans in order to provide arbitrarily small benefits to sentient individuals. Intuitively, that would be monstrous.

P-Vulcans may also seem to challenge the view developed in this paper. If only desires associated with positive affective states count toward well-being, it might seem to follow that p-Vulcans aren't welfare subjects relative to the desire-fulfilment theory. If being a welfare subject is necessary for moral standing, it follows that p-Vulcans lack moral standing.

In response, we can simply deny the assumption that the absence of felt pleasantness or unpleasantness entails the absence of affect. That entailment is likely to strike us as plausible insofar as we assume that the valence property essential to any affective state consists in felt (un)pleasantness. However, there are excellent reasons to reject a hedonic theory of valence (Prinz 2004: 167–168; Carruthers 2017). Most obviously, the possibility of unconscious emotions rules out the hypothesis that hedonic experience is a necessary ingredient in affective states. Moreover, it is plausible that the hedonic quality of an emotional experience can be inverse to the valence of the emotion itself, given that people seem to enjoy horror movies and other forms of art intended to evoke negative emotions (Gaut 1993).

It might be objected that p-Vulcans should not be imagined as simply undergoing experiences that involve all the same physiological, experiential, and cognitive perturbations that we associate with joy or anger except for the absence of felt (un)pleasantness. Rather, we should imagine p-Vulcans as completely devoid of all vestiges of affect (Smithies forthcoming). In that case, however, I feel quite comfortable in relying on the arguments rehearsed in section 3 to reject the claim that p-Vulcans can instantiate the welfare good of desire-fulfilment (compare Heathwood 2019: 680–682).

This need not commit us to the view that p-Vulcans can permissibly be killed without justification. We could instead drop the assumption that desire-fulfilment is the only welfare good. Alternatively, we can drop the assumption – shared by Singer and Chalmers – that welfare sub-

jectivity is necessary for moral standing. It may be that the psychological profile of a p-Vulcan – including the crucial property of phenomenal consciousness – entitles them to some kind of respect, even if they cannot be benefited or harmed (compare Kamm 2007: 228–229; Kagan 2019: 34–35).

6 Conclusion and Implications

The involvement of positive affect appears to be crucial for the desiderative component in the welfare good of desire-fulfilment, which is the one and only good according to desire-fulfilment theories of welfare. Since the view that affective states involve conscious feelings comes naturally to us, it is to be expected that we would assume that the capacity for consciousness is necessary for being a welfare subject if this theory is correct. However, the relationship between affect and consciousness is contested. I think it is reasonable to give at least modest credence to the view that emotional states are not inherently bound to consciousness and that the occurrence of a sentimental mental life would nonetheless suffice for welfare subjectivity and moral status in the absence of consciousness. I think we should be a good deal more confident that the involvement of something like positive emotion is a necessary condition for the desiderative component in the welfare good of desire-fulfilment.

If something like this package of claims is correct, then our moral concern for the welfare of non-human animals need not hinge on the question of whether they have phenomenal consciousness (compare Dawkins 2017; Kagan 2019: 25–27). Granting that affective states aren't intrinsically bound to consciousness, it seems to be an open possibility that some animals might not be conscious but might nonetheless be capable of emotions like fear or anxiety.

For example, we can have only limited confidence that flies are conscious (Birch 2022). Nonetheless, there is evidence indicative of the occurrence of emotion-like states in *Drosophila* (Adolphs and Anderson 2018: 203–210). Repeated presentation of a stimulus simulating the approach of an aerial predator is able to induce persistent, scaleable states of defensive arousal and threat in flies (Gibson et al. 2015). Rates of moving and hopping increased with the number of presentations of the stimulus, and elevated levels of moving and hopping persisted after the stimulus went away, decaying only gradually back to baseline over a period of tens of seconds. The

number of passes made by the overhead stimulus also influenced decisions made by starved flies about whether to remain or return to a food patch. Flies have also been found to be more willing to explore open arenas they otherwise avoid following administration of diazepam, a well-known anxiolytic drug (Mohammad et al. 2016).

On the other hand, the hypothesis that affective states are necessary for the desiderative component in the welfare good of desire-fulfilment may pose obstacles to attributions of welfare to certain kinds of AI systems.

Large language models (LLMs) like OpenAI's ChatGPT have catapulted AI into public consciousness, stimulating renewed interest in questions about what it takes for a digital system to be conscious (Chalmers 2023) or qualify as a welfare subject (Goldstein and Kirk-Giannini ms). Although LLMs have found applications in robotics (Driess et al. 2023; Ahn et al. 2024), they are typically disembodied. They have neither a body to call their own, nor any representation thereof. In this respect, they are profoundly unlike human minds (Chemero 2023).

As noted previously, a range of different theories of emotion converge on the idea that emotions essentially involve interoceptive representations of bodily states. It is not only the James-Lange theory and its modern descendants (James 1884; Lange 1885; Damasio 1994, 1999; Prinz 2004). In Lazarus' development of appraisal theory, the physiological changes associated with emotion are said to be what gives emotions their 'heat' and sets them apart from cold cognitions (Lazarus 1991). In Barrett's constructivist theory of emotion, emotion categories are concepts whose function is to interpret somatic changes registered interoceptively (Barrett 2017).

If affective states generally involve interoceptive representations of bodily states and positive affective states are necessary for attitudinal desire, it follows that AI systems need to engage in something like interoceptive monitoring of apparent internal physical states if they are to instantiate the welfare good of desire-fulfilment (Graziano 2019: 131–132). This would restrict attributions of the welfare good of desire-fulfilment to future AI systems that retain the disembodied character of LLMs like ChatGPT, even if they exhibit markers of phenomenal consciousness (Butlin et al. 2023).

On the other hand reflecting on the case of AI systems might lead us to worry that theories of affect that posit a central role for the interoceptive registrations of bodily states are parochial and should not be extended beyond the case of biological minds. Theories of that kind might account

for the way emotions arise in organisms with brains made of neurons and glia, but perhaps fall short of the level of generality and abstraction needed to account for the novel ways familiar mental states could be realized in unliving things.[14]

ACKNOWLEDGMENTS

For comments on previous drafts of this paper, I'm especially grateful to Patrick Butlin, Mattia Cecchinato, Tomi Francis, Brad Saad, Teru Thomas, and Hayden Wilkinson, as well as to audiences at the AI, Animals, and Digital Minds conference at LSE in June 2024, the London AI and Humanity Project at the Institute of Philosophy in October 2024, and the Hong Kong Catastrophic Risk Centre at Lingnan University in October 2024.

[14]I'm grateful to Jeff Sebo for this suggestion.

References

Adolphs, R. and D. J. Anderson. 2018. *The Neuroscience of Emotion: A New Synthesis*. Princeton, NJ: Princeton University Press.

Ahn, M., D. Dwibedi, C. Finn, M. G. Arenas, K. Gopalakrishnan, K. Hausman, B. Ichter, A. Irpan, N. Joshi, R. Julian, S. Kirmani, I. Leal, E. Lee, S. Levine, Y. Lu, I. Leal, S. Maddineni, K. Rao, D. Sadigh, P. Sanketi, P. Sermanet, Q. Vuong, S. Welker, F. Xia, T. Xiao, P. Xu, S. Xu, and Z. Xu. 2024. AutoRT: Embodied foundation models for large scale orchestration of robotic agents. arXiv:2401.12963. URL: https://arxiv.org/abs/2401.12963.

Arico, A., B. Fiala, R. F. Goldberg, and S. Nichols. 2014. The folk psychology of consciousness. In *Experimental Philosophy: Volume 2*. Oxford: Oxford University Press.

Arneson, R. J. 1999. Human flourishing versus desire satisfaction. *Social Philosophy and Policy 16*(1): 113–142.

Baars, B. J. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Bain, D. T. 2003. Intentionalism and pain. *Philosophical Quarterly 53*(213): 502–523.

Barlassina, L. and M. K. Hayward. 2019. More of me! Less of me! Reflexive imperativism about affective phenomenal character. *Mind 128*(512): 1013–1044.

Barrett, L. F. 2017. *How Emotions Are Made: The Secret Life of the Brain*. London: Macmillan.

Berger, J., B. Fischer, and J. Gottlieb. ms. Minds matter. Unpublished manuscript.

Berridge, K. and P. Winkielman. 2003. What is an unconscious emotion? (The case of unconscious "liking"). *Cognition and Emotion 17*(2): 181–211.

Birch, J. 2022. The search for invertebrate consciousness. *Noûs 56*(1): 133–153.

Birch, J. 2024. *The Edge of Sentience: Risk and Precaution in Humans, Other Animals, and AI*. Oxford: Oxford University Press.

Bourget, D. and A. Mendelovici. 2019. Phenomenal intentionality. In *The Stanford Encyclopedia of Philosophy* (Fall 2019 ed.)., ed. E. N. Zalta. Metaphysics Research Lab, Stanford University.

Bradford, G. 2022. Consciousness and welfare subjectivity. *Noûs 4*(75): 905–921.

Bradley, B. 2014. Objective theories of well-being. In *The Cambridge Companion to Utilitarianism*, eds. B. Eggleston and D. Miller, 220–238. Cambridge: Cambridge University Press.

Brandt, R. B. 1979. *A Theory of the Good and the Right*. Oxford: Oxford University Press.

Browning, H. and J. Birch. 2022. Animal sentience. *Philosophy Compass 17*(5): e12822.

Bush, G. H. W. 1990. *Public Papers of the President of the United States: George Bush, 1990 - July 1 to December 31, 1990*. Washington, D. C.: U.S. Government Publishing Office.

Butlin, P., R. Long, E. Elmoznino, Y. Bengio, J. Birch, A. Constant, G. Deane, S. M. Fleming, C. Frith, X. Ji, R. Kanai, C. Klein, G. Lindsay, M. Michel, L. Mudrik, M. A. K. Peters, E. Schwitzgebel, J. Simon, and R. VanRullen. 2023. Consciousness in artificial intelligence: Insights from the science of consciousness. https://arxiv.org/abs/2308.08708.

Carruthers, P. 2017. Valence and value. *Philosophy and Phenomenological Research 97*(3): 658–680.

Chalmers, D. J. 2022. *Reality +: Virtual Worlds and the Problems of Philosophy*. London: Allen Lane.

Chalmers, D. J. 2023. Could a large language model be conscious? arXiv:2303.07103. URL: https://arxiv.org/abs/2303.07103.

Chemero, A. 2023. LLMs differ from human cognition because they are not embodied. *Nature Human Behaviour* 7: 1828 – 1829.

Clark, A. 1994. Beliefs and desires incorporated. *Journal of Philosophy 91*(8): 404–25.

Copp, D. 2006. On the agency of certain collective entities: An argument from "normative autonomy"'. *Midwest Studies in Philosophy 30*(1): 194–221.

Craig, A. 2003. A new view of pain as a homeostatic emotion. *Trends in Neurosciences 26*(6): 303–307.

Cutter, B. and M. Tye. 2011. Tracking representationalism and the painfulness of pain. *Philosophical Issues 21*(1): 90–109.

Damasio, A. R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: Avon Books.

Damasio, A. R. 1999. *The Feeling of What Happens: Body, Emotion, and the Making of Consciousness*. London: William Heinemann.

Dawkins, M. S. 2017. Animal welfare with and without consciousness. *Journal of Zoology 301*(1): 1–10.

de Lazari-Radek, K. and P. Singer. 2014. *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. Oxford: Oxford University Press.

Dehaene, S., J.-P. Changeux, L. Naccache, J. Sackur, and C. Sergent. 2006. Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences 10*(5): 204–211.

Deonna, J. A. and F. Teroni. 2012. *The Emotions: A Philosophical Introduction*. New York, NY: Routledge.

Driess, D., F. Xia, M. S. M. Sajjadi, C. Lynch, A. Chowdhery, B. Ichter, A. Wahid, J. Tompson, Q. Vuong, T. Yu, W. Huang, Y. Chebotar, P. Sermanet, D. Duckworth, S. Levine, V. Vanhoucke, K. Hausman, M. Toussaint, K. Greff, A. Zeng, I. Mordatch, and P. Florence. 2023. PaLM-E: An embodied multimodal language model. arXiv:2303.03378. URL: https://arxiv.org/abs/2303.03378.

Freud, S. 1915 / 1963. The unconscious. In *The Collected Papers of Sigmund Freud - General Psychological Theory: Papers on Metapsychology*, ed. P. Rieff. New York, NY: Collier Books.

Gaut, B. 1993. The paradox of horror. *British Journal of Aesthetics 33*(4): 333–345.

Gibson, W. T., C. R. Gonzalez, C. Fernandez, L. Ramasamy, T. Tabachnik, R. R. Du, P. D. Felsen, M. R. Maire, P. Perona, and D. J. Anderson. 2015. Behavioral responses to a repetitive visual threat stimulus express a persistent state of defensive arousal in *Drosophila*. *Current Biology 25*(11): 1401–1415.

Gilbert, M. 2002. Collective guilt and collective guilt feelings. *The Journal of Ethics 6*(2): 115–143.

Goldie, P. 2002. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.

Goldstein, S. and C. D. Kirk-Giannini. ms. AI wellbeing. URL: https://philpapers.org/rec/GOLAWE-4.

Graziano, M. S. A. 2019. *Rethinking Consciousness: A Scientific Theory of Subjective Experience*. New York, NY: W.W. Norton & Company.

Hatzimoysis, A. 2007. The case against unconscious emotions. *Analysis 67*(4): 292–299.

Heathwood, C. 2019. Which desires are relevant to well-being? *Noûs 53*(3): 664–688.

Heathwood, C. 2022. Ill-being for desire satisfactionists. *Midwest Studies in Philosophy* 46: 33–54.

Hess, K. M. 2014. Because they can: The basis for the moral obligations of (certain) collectives. *Midwest Studies in Philosophy 38*(1): 203–221.

Hess, K. M. 2018. Does the machine need a ghost? Corporate agents as nonconscious Kantian moral agents. *Journal of the American Philosophical Association 4*(1): 67–86.

Horgan, T. and J. Tienson. 2002. The intentionality of phenomenology and the phenomenology of intentionality. In *Philosophy of Mind: Classical and Contemporary Readings*, ed. D. J. Chalmers, 520–533. Oxford: Oxford University Press.

James, W. 1884. What is an emotion? *Mind 9*(34): 188–205.

Kagan, S. 2014. An introduction to ill-being. *Oxford Studies in Normative Ethics* 4: 261–88.

Kagan, S. 2019. *How to Count Animals, More or Less*. Oxford: Oxford University Press.

Kamm, F. M. 2007. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Oxford: Oxford University Press.

Klein, C. 2007. An imperative theory of pain. *Journal of Philosophy 104*(10): 517–532.

Kraut, R. 1994. Desire and the human good. *Proceedings and Addresses of the American Philosophical Association 68*(2): 315.

Lange, C. 1885. *Om Sindsbevægelser: Et Psyko-Fysiologisk Studie*. Copenhagen: Jacob Lunds Forlag.

Lazarus, R. S. 1991. Progress on a cognitive-motivational-relational theory of emotion. *The American psychologist 46*(8): 819–34.

Lee, A. forthcoming. Consciousness makes things matter. *Philosophers' Imprint*.

Lin, E. 2020a. Attitudinal and phenomenological theories of pleasure. *Philosophy and Phenomenological Research 100*(3): 510–524.

Lin, E. 2020b. The experience requirement on well-being. *Philosophical Studies 178*(3): 867–886.

Lin, E. 2022. Two kinds of desire theory of well-being. *Midwest Studies in Philosophy* 46: 55–86.

List, C. 2021. Group agency and artificial intelligence. *Philosophy and Technology* 34: 1213–1242.

List, C. and P. Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.

Loar, B. 2003. Phenomenal intentionality as the basis of mental content. In *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, eds. M. Hahn and B. Ramberg, 229–258. Cambridge, MA: MIT Press.

Manning, R. C. 1984. Corporate responsibility and corporate personhood. *Journal of Business Ethics 3*(1): 77–84.

Martínez, M. and L. Barlassina. forthcoming. The informational profile of valence: The metasemantic argument for imperativism. *British Journal for the Philosophy of Science*.

Mohammad, F., S. Aryal, J. Ho, J. C. Stewart, N. A. Norman, T. L. Tan, A. Eisaka, and A. Claridge-Chang. 2016. Ancient anxiety pathways influence *Drosophila* defense behaviors. *Current Biology 26*(7): 981–986.

Nussbaum, M. C. 2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.

Öhman, A. and J. Soares. 1994. "Unconscious anxiety": Phobic responses to masked stimuli. *Journal of Abnormal Psychology 103*(2): 231–40.

Overvold, M. C. 1982. Self-interest and getting what you want. In *The Limits of Utilitarianism*, eds. H. B. Miller and W. H. Williams, 186–94. Minneapolis, MN: University of Minnesota Press.

Pallies, D. 2022. Attraction, aversion, and asymmetrical desires. *Ethics 132*(3): 598–620.

Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Pettit, P. 2007. Responsibility incorporated. *Ethics 117*(2): 171–201.

Pitcher, G. 1965. Emotion. *Mind 74*(295): 326–346.

Prinz, J. J. 2004. *Gut Reactions: A Perceptual Theory of the Emotions*. Oxford: Oxford University Press.

Prinz, J. J. 2012. *The Conscious Brain: How Attention Engenders Experience*. Oxford: Oxford University Press.

Rawls, J. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rolls, E. T. 2013. *Emotion and Decision-Making Explained*. Oxford: Oxford University Press.

Rorty, R. 1970. Incorrigibility as the mark of the mental. *The Journal of Philosophy 67*(12): 399–424.

Russell, J. 1980, 12). A circumplex model of affect. *Journal of Personality and Social Psychology* 39: 1161–1178.

Russell, J. A. 2003. Core affect and the psychological construction of emotion. *Psychological Review 110*(1): 145–172.

Scarantino, A. 2010. Insights and blindspots of the cognitivist theory of emotions. *British Journal for the Philosophy of Science 61*(4): 729–768.

Schroeder, T. 2004. *Three Faces of Desire*. New York, US: Oxford University Press.

Searle, J. R. 1979. What is an intentional state? *Mind 88*(349): 74–92.

Searle, J. R. 1990. Consciousness, explanatory inversion and cognitive science. *Behavioral and Brain Sciences 13*(1): 585–642.

Searle, J. R. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.

Silver, K. 2019. Can a corporation be worthy of moral consideration? *Journal of Business Ethics 159*(1): 253–265.

Singer, P. 1993. *Practical Ethics*, 2nd. ed. Cambridge: Cambridge University Press.

Smith, M. 1987. The Humean theory of motivation. *Mind 96*(381): 36–61.

Smithies, D. 2012. The mental lives of zombies. *Philosophical Perspectives 26*(1): 343–372.

Smithies, D. forthcoming. Affective consciousness and moral status. *Oxford Studies in Philosophy of Mind*.

Solomon, R. C. 1973. Emotions and choice. In *Not Passion's Slave*, 3–24. Oxford: Oxford University Press.

Stalnaker, R. C. 1984. *Inquiry*. Cambridge, MA: MIT Press.

Sumner, L. W. 1996. *Welfare, Happiness, and Ethics*. Oxford: Oxford University Press.

Sumner, W. 2020. The worst things in life. *Grazer Philosophische Studien 97*(3): 419–432.

Tye, M. 1995. A representational theory of pains and their phenomenal character. *Philosophical Perspectives* 9: 223–39.

Vadas, M. 1984. Affective and nonaffective desire. *Philosophy and Phenomenological Research 45*(2): 273–80.

van der Deijl, W. 2020. The sentience argument for experientialism about welfare. *Philosophical Studies 178*(1): 187–208.

Whiting, D. 2018. Consciousness and emotion. In *Routledge Handbook of Consciousness*, ed. R. J. Gennaro, 310–321. New York, NY: Routledge.

Winkielman, P. and K. C. Berridge. 2004. Unconscious emotion. *Current Directions in Psychological Science 13*(3): 120–123.

Winkielman, P., K. C. Berridge, and J. L. Wilbarger. 2005. Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin 31*(1): 121–135.

Winkielman, P. and Y. Gogolushko. 2018. Influence of suboptimally and optimally presented affective pictures and words on consumption-related behavior. *Frontiers in Psychology* 8: 2261.

Wringe, B. 2014. May I treat a collective as a mere means? *American Philosophical Quarterly 51*(3): 273–284.