



# The Axiology of Pain and Pleasure

Alycia LaGuardia-LoBianco<sup>1</sup> · Paul Bloomfield<sup>2</sup> 

Accepted: 21 March 2023

© The Author(s), under exclusive licence to Springer Nature B.V. 2023

## 1 Introduction

There is little more commonplace, in English, than using the word “good” as a synonym for “pleasant”. When we eat something we like, we say, “it tastes good”, and when something feels pleasant, we say, “it feels good”. Similarly, “bad” is commonly used as a synonym for “unpleasant” or “painful”. If food is over-salted, we say that “it tastes bad”; if music is cacophonous, “it sounds bad.” It is hard to deny the idea that having a toothache is a “bad feeling”.

From a philosophical point of view, thoroughly uncontroversial claims are hard to come by, but they are most easily found in explications of what is natural and commonplace in the way we speak. One good candidate is the set of claims involved in maintaining the ideas that “pleasure is good” and “pain is bad”, such that “it is pleasurable” entails “it is good” and “it is painful” entails “it is bad”. Commonsense gets crystallized as philosophy via the concept of *intrinsic value*, such that pleasure is thought to be intrinsically valuable or good in itself, and pain is similarly intrinsically disvaluable or bad in itself.

The “ordinary view” of pleasure and pain is loath to accept that there could be anything morally bad about “pleasure in itself”: people love their pleasure. Commonsense says that while pleasures may bring about prudentially bad consequences (disease, poverty) and pains may bring about prudentially good consequences (victory, a newborn infant), there is never anything inherently (morally) bad about pleasure or inherently (morally) good about pain; pleasure is always good because its value is in itself, and pain always bad because its disvalue is in itself, though these things may bring about other kinds of values. Analogous thoughts attach to those causes of pleasure which are themselves not good (malice, revenge): these do not affect the positive value of the pleasurable sensations they engender. Good things

---

✉ Paul Bloomfield  
phsb@uconn.edu

Alycia LaGuardia-LoBianco  
laguarda@gvsu.edu

<sup>1</sup> Philosophy Department, Grand Valley State University, Allendale, MI, USA

<sup>2</sup> Department of Philosophy, UConn, 344 Mansfield Rd, U-1054, Storrs, CT 06269, USA

are seen as good only “because they result in pleasure and in the relief and avoidance of pain” (*Protagoras*, 354b); pleasures and pains retain these values, everything else notwithstanding.

Following ordinary thought, philosophers have often taken pleasure to be the paradigmatic example of something which is intrinsically good while pain is thought of as archetypically, intrinsically bad. While it is possible to draw up extensive lists of supposed “intrinsic goods” (cf. Frankena 1973) and “intrinsic bads”, it is natural to think that if anything at all is to be found on the lists, pleasure and pain will be at the top. Indeed, as William Frankena presents it, the traditional debate about intrinsic value assumes that pleasure is intrinsically good and the only question is whether anything else is (p. 85). Other philosophers back this up. Irwin Goldstein (1989) puts it like this:

Pleasure is good as such, because of its pleasurable-ness, not because of some further good quality which colors pleasure and may or may not be present. The foundation of pleasure’s goodness, its pleasurable-ness, marks every pleasure. How could pleasure fail to be intrinsically good? (p. 273)<sup>1</sup>

Thus, on commonsense and many philosophical views, there appears to be no distance between the appearance of pleasure or pain and the reality of its positive or negative intrinsic value. Little could seem more obvious, little so (seemingly) carved into nature itself.

Yet entrenched folk theory must sometimes be unearthed: we used to think that species were eternal and Earth was the center of the universe; we used to think space was Euclidian, etc. Our present goal is to dislodge these old and entrenched ideas about pleasure and pain and argue that not only are they false but they are pernicious: there is nothing *intrinsically* good about pleasure nor *intrinsically* bad about pain, and not comprehending this has led to mistakes in practical rationality and morality. We do not deny that, barring rare pathological conditions, pain hurts and pleasure is pleasant, nor do we deny that pain can be bad nor pleasure good. Nor do we deny the existence of intrinsic value in the world, as we think that, at least, virtue and good will have it. Rather, we deny only the entailment of *goodness* from *pleasure* and *badness* from *pain*; we deny only claims that pleasure is intrinsically good and pain is intrinsically bad. Thus, we are giving a classic “error theory” of the Humean sort. Analogous to Hume’s argument that the association between events leads us to mistakenly attribute necessary causes linking them, we argue that the

<sup>1</sup> Other examples: Thomas Hurka (1998) makes the intrinsic goodness of pleasure and the intrinsic badness or “evil” of pain axiomatic in a “general recursive theory of the intrinsic values of attitudes shared by many writers on intrinsic value, from Brentano, Moore, and Ross to Nozick, Chisolm, and Lemos” (p. 316). And while not explicitly accepting a recursive theory like Hurka’s, we can add to the list of those who accept the intrinsic value of pain and pleasure ancient hedonists like Epicurus (1994) and the Cyreniacs (see Annas 1993), as well as modern and contemporary hedonists, utilitarians, and consequentialists like Henry Sidgwick (1981), Peter Singer (1975), Shelly Kagan (1998), Fred Feldman (2004), Alastair Norcross (2004), and Gwen Bradford (2020). Even Kantians are in the game, as can be seen in the work of Christine Korsgaard (1983). Both consequentialist and deontologists have accepted the intrinsic value of pain and pleasure.

consistent association of pleasure and goodness and of pain and badness, respectively, has led to the erroneous attribution of intrinsic value to these sensations. The error we are exposing is axiological, as it does not concern the existence of the goodness of pleasure nor the badness of pain, but only concerns the axiological status of those values.

As will be discussed further below, we understand “intrinsic value” in terms of a value something has in itself as contrasted with extrinsic value or the value something has from a source other than itself. Following Christine Korsgaard (1983), this distinction is conceptually distinct from that of the way in which something is valued as either a final end or as a means. While conceptually distinct, final values and intrinsic values are related. Finlay (2014) writes that “speech and thought about what is ‘good for its own sake’ is indeed about a property of value that is noninstrumental, intrinsic, and necessary... (205)”. Final values depend on intrinsic features of the thing valued; that which we value for its own sake is valued in part because of what it is in itself. While our primary aim is to contest the intrinsic values of pain and pleasure, if we are right, our argument will entail that pain and pleasure are not final ends and are only ever instrumentally (and thus extrinsically) valuable.<sup>2</sup> Additionally, we do not suggest a value monism in our argument. There are different types of value (prudential, moral, aesthetic, etc.), but we will focus on motivating the idea that the *moral* value of pains and pleasures are extrinsic. Ultimately, though, we hold that our argument applies for any possible value of pain and pleasure, whether moral, prudential, etc. (We return briefly to the distinction between moral and prudential value at the end of Section III).

Our argument proceeds by assessing various cases in which pains seem to be good and pleasures seem to be bad. We argue that these cases show that the moral values of pain and pleasure are dependent upon the context in which they are experienced and thus only ever instrumentally valuable. Quite often, the moral value of pains and pleasures line up with their prudential value; quite often, pain is (morally and prudentially) a bad thing given the context in which it is felt, and pleasure is (morally and prudentially) a good thing given its context. But we ought not conclude from these harmonious examples that pleasure is therefore valuable in itself and pain is therefore disvaluable in itself; rather, we suggest that any moral goodness of pleasure and moral badness of pain is attributable to the conditions under which they occur, and thus is instrumental. This suggestion does away with the theoretical need to maintain the intrinsic values of pain and pleasure in the face of counterexamples. Rather than cling to the intrinsic values of pain and pleasure even when it seems clear that some pains are morally good and some pleasures morally bad, we conclude that these values should not be evaluated in isolation but in terms of the context in which the phenomena are experienced.

As Hume argued positively that we should understand causation as “constant conjunction”, we too have a positive argument following the negative conclusion

<sup>2</sup> We take it that if something has instrumental value, it is also extrinsically valuable, but not the converse. So, we will ultimately be arguing for the instrumentality of pain’s badness and pleasure’s goodness, which, if correct, entails its extrinsic value.

that pain is not intrinsically bad and pleasure is not intrinsically good. Specifically, our positive claim is that pleasures and pains are phenomena which are the result of evolution by natural selection and the only value they possess is the (instrumental) value that process gives them. On this view, the moral goodness of pleasure depends on whether it leads us to value what is good and to pursue it, and similarly, pain is good when it helps us avoid what is bad or helps us attend to what ought to be attended to (like an injury). Pleasure is morally bad when it leads us to pursue or do what is bad or wrong, and pain's moral badness depends on whether it keeps us from doing what is good or right. Thus, the values of pleasure and pain are not intrinsic, but instrumental.

As with any error theory, first we will expose the reasons behind the error (Section II), then we will explain what happens as the error occurs (Section III), after which we will explain its prevalence and our positive argument (Section IV). In the conclusion (Section V), we briefly discuss the morally and ethically pernicious results of the error. For the sake of simplicity and due to considerations of scope, we will focus most of our attention on the intrinsic badness of pain, though we take all the basic arguments to apply, *mutatis mutandis*, to pleasure and goodness. We refer to "pains" broadly as including both physical and psychological forms of suffering.

## 2 The Petrification of the Error: On Pain's Supposed Intrinsic Badness

Since philosophers typically take for granted the claim that pain is intrinsically bad and that it therefore ought to be relieved or prevented, there is not much argumentation for what is thought to be so obvious. Philosophically, the clearest explanation of pain's intrinsic disvalue is in what Wayne Sumner (2020) calls "experiential theories of well-being", notably hedonism. On such views, "a subject's well-being [consists] in the felt quality (and quantity) of her affective experience" (Sumner 2020, p. 421). Sumner argues that any theory of well-being ought to account for ill-being, or the ways in which a life can go badly, and that experiential theories can do so by appealing to the presence of pain in a life (2020, p. 421). For hedonism, the "pleasure/pain polarity seems to map in an easy way onto the good/bad polarity for lives (or parts of lives)" (2020, p. 421).<sup>3</sup> Hedonism, of course, relies on the intrinsic badness of pain, and gives an explanation for this in terms of experience: pain feels bad. Classical utilitarianism (particularly Bentham and Mill) represents the commitments entailed by the view that pain is intrinsically bad, but many other non-utilitarian philosophers have taken on this axiological assumption that pain is bad in itself.

<sup>3</sup> For other accounts of ill-being, accounting for "substantive bads" (including pain) is less straightforward. Desire satisfaction theories struggle to account for substantive bads that are not merely privations from unmet desires, or to explain why certain unmet desires are significant enough to be experienced as substantive bads rather than privations (as this would appeal to experiential explanations closed to the desire theorist) (Sumner 2020, pp. 424–427). Additionally, Sumner argues that Richard Kraut's (2007) account of pain as a bad under an objective theory does not acknowledge that sensory pains are often the result of the body functioning *as it should*, and thus cannot account for the badness of pains in this way (Sumner 2020, p. 430). Our positive account below will expand on this last point.

And, as noted (cf. Frankena 1973), others accept the intrinsic value of pleasure and pain though they are not hedonists since they think these are not the only sources of intrinsic value in the world.

The supposed obviousness of the axiological claims is perhaps most evident in Mill's utilitarianism, which is based on the hedonistic assumption that "that pleasure and freedom from pain are the only things desirable as ends" (2001, p. 7). Mill even claims that these foundational assumptions do not admit of standard argument:

Questions of ultimate ends are not amenable to direct proof. Whatever can be proved to be good must be so by being shown to be a means to something admitted to be good without proof...[for example] [t]he art of music is good, for the reason, among others, that it produces pleasure; but what proof is it possible to give that pleasure is good? (p. 4.)

As noted above, though there is a conceptual distinction between intrinsic value and final value, there is a relationship between the two such that what is valuable as a final end will have to depend on why it is valuable for its own sake. Though Mill focuses on goodness as a final end at the beginning of this passage, he seems to suggest that this value depends on what pleasure is – its goodness depends simply on what it is. So, Mill arguably illustrates this relationship. For our purposes, it suggests that the close philosophical association between seeking pleasure as a final end, for its own sake, and the source of that value being the intrinsic features of pleasure. So, here Mill's spade turns, and the same presumably holds for the badness of pain. A theory built upon the intrinsic goodness of pleasure and badness of pain can be accepted or rejected, but Mill does not think the foundational assumptions can be defended by argument.

Similarly, Peter Singer builds his inclusive utilitarian ethics upon the assumption that "Pain and suffering are bad and should be prevented or minimized, irrespective of the race, sex, or species of the being that suffers" (1975, p. 19).<sup>4</sup> The explanation he gives for these assumptions is tautological: all pain ought to be relieved "since pain is pain" (Singer 1975, p. 23). The controversial part of Singer's claim has been his refutation of speciesism, while his assumptions about pain and suffering have not raised an eyebrow. Echoing these claims, political scientist Jamie Mayerfeld equates suffering with the hedonistic notion of pain as a particular negative feeling (1999, p. 16). Like Mill, Mayerfeld thinks our intuitive understanding of suffering as bad does not need proof:

What I claim is that suffering is intrinsically evil for the individual who experiences it—evil in itself. I'm afraid I have no argument for this claim. This doesn't strike me as a great embarrassment, since few moral claims appear to

<sup>4</sup> At places, Singer seems to reserve the term 'pain' for physical pain and 'suffering' for the mental anguish (1975, p. 14), while at others he uses them interchangeably, as when he claims, "Normal adult humans beings have mental capacities which will, in certain circumstances, lead them to suffer more than animals would in the same circumstances" (Singer 1975, p. 17). The distinction does not much matter for our argument, since pain and suffering are both considered to be intrinsically bad whether or not they denote different phenomena.

me more certain. Rather than argue for this claim, I shall simply assert it, and hope that enough of my readers agree (1999 p. 85).

And regarding the claim that that suffering is intrinsically bad from an impersonal perspective, he writes, “I don’t know what I could say in support of a claim that seems so self-evidently true” (1999, p. 87).<sup>5</sup>

Korsgaard (1983) gives utilitarianism an argument to accept the intrinsic badness of pain. In defending a conditional theory of intrinsic value (upon which more below), she writes:

In the early years of this century there was much discussion of the question whether or not a good thing has its value as a result of something like the interest taken in it or the desire someone has for it... Probably the interest in the issue was aroused by a common utilitarian argument that pleasure is the only thing that is good in itself because it is the only thing that we can desire for its own sake (p. 173).

Korsgaard is here setting up an argument for the conclusion that pleasures are intrinsically good when they are conditioned properly by a good will. While what she offers does not constitute a “direct proof” which Mill rightly claims is not forthcoming, she does give a utilitarian reason for thinking that pleasures are intrinsically good and, presumably pains are intrinsically bad: because we can desire to experience pleasures or avoid pains “for their own sake”-- a reason that again relies on the relationship between final and intrinsic value.

Beyond the idea that pleasures are sought and pains avoided “for their own sake”, and thus entail intrinsic (dis)value, or that of locating the badness of pain in how it feels, a further reason that the assumption of pain’s intrinsic badness has been philosophically tenacious is because it explains certain ethical intuitions: if pain is intrinsically bad, then this explains why people are typically motivated to and ought to relieve or prevent pain. This conditional is both powerful and elegant enough to build a moral theory upon, as Bentham, Mill, and others have variously attempted. Considering its place in moral commonsense, it is unsurprising, then, that the assumption that “pain is intrinsically bad” has enjoyed much security and no opposition.

Crucially, both Singer and Mayerfeld hold that the intrinsic badness of pain is sufficient to motivate the moral imperative to relieve or prevent it. Further, both claim that any additional considerations are at best irrelevant and at worst a distraction from the moral assessment of suffering. This is a central part of Singer’s (1975) inclusive ethics since “[t]he evil of pain is, in itself, unaffected by the other characteristics [like the species] of the being that feels the pain” (p. 24). Pain does the moral work all by itself, regardless of who (or what) feels that pain. For Mayerfeld, as for perhaps almost everyone, the very nature of suffering has unequivocal moral

<sup>5</sup> Suffering is typically taken to be a broader category of pain, where pain refers to physically unpleasant sensations. Mayerfeld himself distinguishes between suffering and pain. We draw attention to the method of argumentation, as similar claims have been made about pain.

implications: “We have a prima facie duty to relieve suffering, because suffering is bad and ought not to occur” (1999, p. 111).

The attribution of intrinsic disvalue to pain is motivated by both intuition and moral argument. Yet, we contend that this is an error, and that this error arises from the way in which moral assessment and suffering or pain have been philosophically linked. As shown above, the common cases in which pain’s disvalue is identified are those in which pain is brought about by moral evil or harm. Under this light, the association between pain and its intrinsic disvalue is understandable; why wouldn’t we think that, e.g. the infliction of gross harm and the resulting pain is a bad thing? From here it is a quick step to say that the reason for this intuition is just because of the pain at issue; it is just because of what pain is that it is bad. But because there is already a moral assumption built into these cases – namely, the need to identify the source of the badness at issue – the brunt of the explanation of the badness falls onto the presence of pain. For instance, Singer is starting from well-known examples of harm that also involve pain, and concluding from this that pains are intrinsically bad. Our concern is that the moral tone of these evaluations colors the value assessment of pain rather than isolating it; intrinsic disvalue is attributed to pain before a moral assessment of the circumstances of that pain may account for the disvalue. In other words, it is precisely because pain’s intrinsic disvalue has been found in examples that *already* evince moral badness that the identification of it to pain’s disvalue arises in the first place – that is the source of the error we are pointing to.<sup>6</sup> However, we argue that this initially plausible proposal will not withstand an examination of cases in which pain and badness come apart (unlike those just discussed). Rather, we will suggest that a better theoretical approach to these cases is to deny the intrinsic disvalue of pain.

### 3 Exposing the Error

In order to argue against the intrinsicity of pain’s badness, we propose to look at cases in which, for various reasons, pain’s value does not seem bad, since we take these cases to be counterexamples to the claim that pain is intrinsically bad—the sort of cases which the theories discussed above seek to accommodate—while also providing evidence of the contingency and entailed instrumentality of pain’s badness. These purported counterexamples of pain’s desirability or appropriateness will put pressure on the plausibility of pain’s intrinsic badness. As Korsgaard (1983) notes:

Since intrinsically good things...are thought to have their value *in* themselves, they are thought to have their goodness in any and all circumstances – to carry it with them so to speak. If you find that a certain kind of thing is not good

---

<sup>6</sup> We do not deny that, just as there may be multiple sources of intrinsic value, there may be multiple sources of intrinsic disvalue (Bradford 2021). So, our argument is not that there are *no* intrinsic (dis)values, but just that pain is not one of (possibly) many intrinsically disvaluable things.

in any and all circumstances, that it is good in some cases and not others, its goodness is extrinsic – it is derived from or dependent upon the circumstances (p. 171).

Korsgaard's idea here matches our own conception of *intrinsic value*, which has roots in Plato and Kant. Of course, Plato did not have the term “intrinsic value”, but we think he captures the idea in the exchange between Socrates and Cephalus in *Republic* (330d1–331d). The latter suggests that the virtue of justice might be understood in terms of “returning what is borrowed”, but the former argues against that suggestion by appealing to a situation in which returning what is borrowed is the wrong thing to do. A trait must be good in independent of context in order for it to count as a “virtue”, and we think this is sufficient for saying that virtue is “intrinsically valuable”. Kant's view of intrinsic value seems to be the same, as he opens the first section of his *Groundwork of the Metaphysics of Morals* with the claim that nothing can be “considered good without limitation except a good will” (4:393, 1997). His reasoning is similar to Plato's as he argues that traits like courage are not “good without limitation” because however “undoubtedly good and desirable” they may be “for many purposes” (4:393), “the coolness of a scoundrel makes him not only far more dangerous but also more immediately abominable in our eyes than we would have taken him to be without it” (4:394).

This Platonic/Kantian characterization of intrinsic value is generic and, unsurprisingly, the literature on intrinsic value, on one hand, and the nature of pain (and pleasure), on the other, are deep. In particular, we find the dialectical situation surrounding the claim that “pain is intrinsically bad” complicated by the fact that there are two main theories of pain (and pleasure) – the “felt-quality” and “attitudinal” theories – and there are two main theories of intrinsic value – the Moorean and “conditional” theories. We address these distinctions in turn. Before turning to problematic cases, here are the glosses on the theories themselves.

The felt-quality view says that pain is identical to the felt quality of hurtfulness or unpleasantness, while the attitudinal theories say that one must have some kind of attitude, e.g., dislike or aversion, toward these felt-qualities for there to be pain. If we can call pains and pleasures “items”, then the Moorean view of intrinsic value says that intrinsic values necessarily supervene on the intrinsic nature of the item possessing the value such that this value remains even when the item is considered in isolation from everything else (Bradley 2002). The conditional view of intrinsic value rejects the necessity which the Moorean accepts, claiming that the intrinsic nature of the item per se is not sufficient to determine its intrinsic value and that the item must be placed or considered in certain conditions since, for example, events external to a person's life (after the person dies) can affect the intrinsic value of the person's life (Hurka 1998). All of the philosophers defending these different views agree that pain is intrinsically bad and disagree only on how to best understand this.

The felt-quality theory of pain and the Moorean theory of intrinsic value are the more straightforward options, but these theories face various apparent counterexamples (some to be discussed below) which put pressure on the intuition that pain is indeed intrinsically bad. In response, attitudinal theories of pain and conditional theories of intrinsic value have been developed, seemingly independently, to solve the



problem. (In other words, these theories of pain and intrinsic value were developed in the first place because the more straightforward views end up having counterintuitive results.) One can try to account for pains which are intrinsically bad but good overall by appeal to a Moorean theory of intrinsic value and an attitudinal theory of pain (Bradley 2002), or by appeal to a conditional theory of intrinsic value and a felt-quality view of pain (Kagan 1998). If one adopts a Moorean view of intrinsic value, then the badness of the felt quality of a hurtful experience is considered as a fine-grained state of affairs which can be outweighed (or trumped or silenced) within a more coarse-grained state of affairs involving that negative felt experience and other positive attitudes taken toward it. If one adopts a conditional theory of intrinsic value, then the felt quality of a pain, by itself, has no intrinsic value and the pain becomes intrinsically bad when it is properly “conditioned” by other mental states or circumstances.

Our error theory is supposed to cut across these theories of pain and theories of intrinsic value: whether pain (pleasure) is understood as a felt quality or in terms of certain attitudes, and whether intrinsic value supervenes on intrinsic qualities and/or extrinsic conditions, it is nevertheless a mistake to think that pains are intrinsically bad (and pleasures intrinsically good). That is, we deny that there is an intrinsic disvalue to the felt quality of pain or to the attitudinal states or other conditions associated with it. Preserving the idea that *pain is intrinsically bad* has been a main motivation behind the development of all this philosophical machinery and we suspect that much of the literature on pain/pleasure and intrinsic value commits this error.

So, we intend to show that pain is not intrinsically bad by pointing to circumstances which show, borrowing from Korsgaard, that it is “not [bad] in any and all circumstances, that it is [bad] in some cases and not in others”. We begin with morally loaded examples, such as the way we ought to feel pain when grieving the loss of a loved one, witnessing injustice, or empathizing with a suffering patient. After this, we consider harmlessly enjoyable cases of what is known as “benign masochism” and other similar, though non-benign, pleasures. Further below, to a lesser extent, we consider cases of “bad pleasures”.

These first pains are not harmlessly enjoyable, but rather are valuable because it is good and right to feel them. The pain of grief, say, seems to be good since it is what happens to love when the beloved dies; not grieving for someone at all implies not loving that person at all. As Robert Solomon argues, “[Grief is] the continuation of love,” (2004, p. 90), it “is not only expected, as the *appropriate* reaction to the loss of a loved one, but also in a strong sense is *obligatory*” (2004, p. 75). To claim that even this pain is intrinsically bad would misunderstand an important human experience and its place in our lives. No one denies that grief is painful. But the inference from grief’s painfulness to grief’s intrinsic badness is fallacious. Here is a clear example of the error we wish to point out: however unpleasant it may be, upon the loss of a loved one, it is (morally) good to grieve (appropriately), and this puts pressure on the claim that it is intrinsically bad. Grief is an unpleasant healing process but there is nothing to be gained by insisting that there is something intrinsically bad about it, however unpleasant it is: if there were something intrinsically bad about grief, this would be a reason to avoid it, and yet it is morally dubious to think unpleasantness is a reason to avoid grieving over the death of a loved one.

So, while one may arrange any number of theoretical maneuvers to accommodate this intuition while maintaining pain's intrinsic badness, we suggest that the more straightforward and appropriate interpretation of this and similar cases (witnessing injustice, sympathizing with a suffering patient) is that pain, however unpleasant, is not intrinsically bad, and not bad at all in some cases.

Consider now cases of "benign masochism," in which people report enjoying difficult, painful, unpleasant, or discomfiting experiences like the burn of spicy food, the sadness after reading a literary tragedy, or the fear of thrill-rides like roller coasters (Rozin et al. 2013, p. 439).<sup>7</sup> These sorts of pains can be desirable for their own sake, and thus do not appear to be intrinsically bad.

A further set of related counterexamples derives from a non-academic source amidst a community of practitioners of a sport: consider a distinction used widely among rock climbers and alpinists (Samet 2011, p. 100).<sup>8</sup> Alpine climbing involves technical and difficult ascents of rock, snow, and/or ice routes in mountains, where the point of the activity is the climb itself rather than reaching the summit. (Reaching the summit is the point of mountaineering.) The relevant distinction is between Type 1 Fun and Type 2 Fun, and was first pointed out around 1985 by Rainer Newberry, a geology professor at the University of Alaska, to mark a "fun-to-suffering" ratio. Type 1 Fun is thought of as "true fun" and understood as "enjoyable while it is happening". Examples include "good food, good sex, 5.8 hand cracks, sport climbing, powder skiing, margaritas" (p. 85).<sup>9</sup> Type 2 Fun, on the other hand, is "fun only in retrospect, hateful [it hurts or is unpleasant] while it's happening." Examples are "working out till you puke, and usually ice and alpine climbing". Importantly, alpine climbing is not an example of benign masochism because it is truly dangerous (Rozin et al., p. 440).

The point for our purposes is that the pain involved in Type 2 Fun is not thought of as "bad" by alpinists but, quite to the contrary, as something to be sought as a particular kind of fun. This sort of fun is, of course, not had only by alpinists but generalizes to anyone who enjoys any kind of hard challenges, including non-physical challenges like writing a philosophical dissertation, or even solving a difficult crossword puzzle. These are activities which are unpleasantly vexing and frustrating yet this is a desired aspect of the experience: there is little enjoyment in completing

<sup>7</sup> Rozin et al. list twenty-nine examples of such benign masochism: sad movies, sad novels, crying to sad movies, sad music, sad painting, burn mouth, spicy food, eyes tearing, tacos with hot sauce, sweating, disgusting jokes, disgusting experiences, pinching pimples, picking nose, pounding heart, frightening movies, thrill rides, massage pain, flashes cold pain, flashes hot pain, beer, scotch, physically exhausted, physically active, bitter foods, unsweetened coffee, gory movies, stinky cheese, and anger fiction (p. 441).

<sup>8</sup> We thank Philip Ebert for informing us of this distinction.

<sup>9</sup> The authors had to ask Ebert what 5.8 hand cracks are and why they are so fun and enjoyable. He responded as follows, "A 5.8 hand crack is fairly straightforward and tends to be an enjoyable climb at least for more experienced climbers (there is some elitism in that definition). First off, 5.8 is not that hard (the 5.8 is a grade that goes up to 5.15); a 'hand-crack means that it's a climb that is a crack where your whole hand fits in (and not merely a finger), so from a technical perspective it is not that demanding and there is a low risk of injury. Finally, hand-cracks are really easily protectable, and given your whole hand is stuck in a crack, it's rare to fall and if you do you won't die or even get hurt."

simple crossword puzzles. It is the unpleasantness of the hard challenge, of being tested when one knows failure is a real possibility, which makes the activity enjoyable. Were there no displeasure, difficulty, frustration, discomfort, or pain involved, there would be no hard challenge—and thus, no Type 2 Fun.<sup>10</sup>

It is not that the difficulty of a hard challenge is experienced as pleasurable (working out till one vomits is not pleasurable; neither is being frustrated by a crossword puzzle clue), nor is there any reason to think that some hedonic calculation is necessarily involved in which the supposedly intrinsic disvalue of the pain is outweighed by the goodness of the pursuit as a whole. Evidence for this is that when a pursuit becomes easy, lacking its unpleasantness, it is typically left behind for a new, hard challenge.

Being challenged is necessarily unpleasant and there is nothing bad about this unpleasantness in the least, the evidence for this being that we seek out such challenges when we are at leisure activities: being challenged is good (Type 2) fun.

To assess these claims given the theories under consideration, we begin with the felt quality view of pain and then attend to the attitudinal view. The felt quality view of pain has trouble making sense of the goodness of the kinds of pain discussed above. On this view, the badness of the unpleasantness of Type 2 Fun must be weighed against (or somehow measured against) the goodness of the pleasure of the overall experience. The problem with such a move is that it yields the wrong results. Here we employ an argument of Gwen Bradford's (2020) called the "zero value objection". In her discussion of what she calls "hurts so good" experiences, which are similar to examples of benign masochism and Type 2 Fun, Bradford rightly points out that if we simply engaged in a hedonic calculation of the intrinsic unpleasantness of, say, Type 2 Fun with the pleasure of the experience of the whole, then the pleasure and displeasure would cancel each other out and the experience would have zero hedonic value, which it obviously does not. And the problem becomes more significant when we realize that the hotter the chili pepper or the more difficult the alpine climb, the more displeasure should be added into the hedonic calculation, and this should degrade the overall pleasure of the experience when in fact it does not and can even have the opposite effect.

It is more apt to understand the people who pursue hard challenges as doing so because of their difficulty, which is exactly what is valuable about Type 2 Fun and hard challenges: good experiences to have, however unpleasant they may be. Hard challenges can only be valuable and good not *in spite* of their unpleasantness but rather *because* of it. Were this unpleasantness intrinsically bad, it is hard to understand how experiencing something intrinsically bad, and doing so for the sake of having fun, can be valuable and good without there being a mistake somewhere. But it is odd to think all alpinists are somehow mistaking badness for goodness. If all unpleasantness is intrinsically bad, it seems impossible to explain why so many

<sup>10</sup> The idea of experiencing "flow" requires this sort of challenge. Nakamura and Csikszentmihalyi write that the experience of flow is the result of "engaging just-manageable challenges by tackling a series of goals, continuously processing feedback about progress, and adjusting action based on this feedback" (2009, p. 90).

people would voluntarily experience something intrinsically bad, doing so “for fun”, when often it is the case that more (supposed) intrinsic badness there is, the greater the challenge there is, and so the greater the fun and the better the experience. Again, though these cases could be explained away while maintaining pain’s intrinsic value, we see that doing so is in tension with common experiences and elegant explanations. Denying pain’s intrinsic badness is the more plausible and theoretically attractive option when compared to the felt quality theory of value.

Notice how the present discussion of Type 2 Fun and the zero-value objection would also apply to the enjoyment of a hard challenge that is immoral. Imagine an assassin taking pleasure in making a long sniper’s shot, enjoying the difficulty of the challenge the way an alpinist enjoys a challenging climb. The assassin’s act is morally worse because of the pleasure, not in spite of it. This example shows commonsense to have counterintuitive axiological implications. On one side of the folk view, pain is intrinsically bad and pleasure is intrinsically good, yet commonsense morality also would say that the assassin’s pleasure makes the act worse overall. If pleasure is intrinsically good, then adding it to what the assassin does should improve the overall value of the act, not make it worse. But it does make it worse.

The point regarding the bad pleasure of the assassin is fortified by another case discussed by Valerie Tiberius (2018). It is unfortunately horrific and, like discussions involving the Holocaust, it is important to not trivialize the horror while considering it as a philosophical example; indeed, the example’s power comes from its horror. The example is that some victims of incest, rape, or sexual assault experience pleasure while they are being violated. It seems likely that everyone not wedded to the idea that pleasure is intrinsically valuable will repudiate any supposed goodness in this involuntary pleasure —indeed, like the assassin’s sniper shot, the fact that a survivor of sexual assault felt any pleasure at all makes the assault that much worse. At best, those who defend the intrinsic goodness of pleasure will have to acknowledge that this case represents what would normally be seen, in non-philosophical contexts, as a poisonous bullet which simply cannot be bitten. This case alone should cast serious doubt on the supposed intrinsic goodness of pleasure. (Indeed, consideration of the example itself is so horrific that we resist using it casually in the course of argument, though it does come up one more time at the end of this section.)

The attitudinal theories of pleasure have more resources to deploy in discussions of these cases. For example, Feldman’s (2004) view of pleasure and pain is that pleasure occurs when people take an attitude of enjoyment toward some state of affairs or are glad the state of affairs obtains. Displeasure or pain occurs when someone is displeased by some state of affairs. So, when alpinists enjoy their Type 2 Fun, they take attitudinal pleasure in what they are feeling, so this counts as pleasure regardless of the felt quality of the experience. Problems arise, however, when we consider again the assassin’s pleasure at making the long sniper shot or, as Noah Lemos puts it, “What is it that makes the wicked man’s being happy...so offensive?” (1994, p. 43)

And here attitudinal theories encounter trouble. In order to account for our condemnation of the assassin’s pleasure or the wicked man’s happiness, Feldman claims that attitudinal pleasures and pains must be “adjusted” by appropriateness

or conditions of desert (2004, p. 119). So, the assassin's pleasure is not intrinsically good because it is not deserved. This is to conjoin the attitudinal theory of pain (pleasure) with the conditional theory of intrinsic value. Bradford's (2020) "reverse conditionalism" makes an analogous move. Where Feldman uses desert to adjust the assassin's pleasure to keep it from being intrinsically good, Bradford claims that felt qualities of painful experiences are intrinsically bad *unless* there is a relevant positive attitude which "defeats" its badness, as in cases of "hurts so good" experiences (p. 247).<sup>11</sup> Both these moves are meant to accommodate our intuitions that there is nothing intrinsically good about the assassin's pleasure nor is there anything intrinsically bad about Type 2 Fun or "hurts so good" experiences.

Metaphysical problems arise, however, when we notice that claims about the existence of intrinsic value require making ontological posits: we are saying something exists. Indeed, the value exists "intrinsically", whatever we take this word to mean. Commonsense says there is some relation between pains and pleasures occurring and goodness or badness coming into existence. These "intrinsic values" take "the value of bound variables" in a normative, axiological theory. Sticking with pains, if pains are intrinsically bad, then when a pain comes into existence, badness comes into existence too. And reflection on this keeps the maneuvers within Feldman's and Bradford's theories from doing their work. For Feldman: how do circumstances of desert *adjust* the assassin's pleasure in such a way that it prevents an intrinsic value from coming into existence or reversing its "hedonic polarity"? For Bradford: how does an attitude *defeat* an intrinsic value? In a footnote, Bradford explains "defeat" in terms of "change": certain attitudes, which are a part of the whole, *change* the intrinsic value of pain which is another part of the whole (p. 248, note 16). But from an ontological point of view, Feldman's "adjustment" and Bradford's "defeat" are equally mysterious. Both would have to claim that quantities of intrinsic value only come into being after possible adjustment or defeat, but as processes *adjustment* and *defeat* ontologically presuppose something being adjusted or defeated. If the assassin's pleasure is adjusted to be bad, it must first be good; if it is defeated, it must at least first exist.

Claiming that pleasures and pains are "complex states of affairs", as Feldman does (p. 120), does not mitigate the mystery. He writes, "The theory will then say that the intrinsic value of an episode of pleasure is equal to the desert-adjusted amount of pleasure it contains, rather than...the *raw* amount of pleasure it contains" (p. 120, italics in original). Notice that this presupposes some normal proportional

<sup>11</sup> Bradford (2020) uses the term "non-instrumental" in place of "intrinsic", but this causes terminological problems: replacing the latter with the former leads to a conflation of intrinsic value and constitutive value. "Constitutive value" marks the presence of the part/whole relation, which is distinct from intrinsic value. Take the pain involved in an injury, say a cracked rib, and let's assume the entire sensation can be broken up, to keep it simple, into the acute pain in the rib and the black and blue ache of the surrounding area. The ache contributes disvalue to the sensation, but it is not instrumental to it nor does it seem right to say that it is intrinsic to the sensation. Rather, we want to say that the ache is part of the sensation -- it partly constitutes the sensation, as my hand partly constitutes my body. Saying that a pleasure is something we can desire "for its own sake" is better, as long as it does not covertly imply the reification of sakes, though it may. The worry is that without reification, the only alternative is to say that when we want something "for its own sake" we are saying that we want it for no reason at all.

relationship between how much raw pleasure is experienced and how much intrinsic value it has. How do circumstances affect how much intrinsic value is generated by a certain amount of raw pleasure if the value is intrinsic to the pleasure? While these adjustments or defeats are needed to prevent these theories from contradicting our intuitions about the sniper's pleasure and the wicked man's happiness, they have, as Russell would put it, all the benefits of theft over honest toil: the challenge for Feldman is that even if we assume that pains are attitudinal, *how* do circumstances of desert adjust the intrinsic value of a pain prior to the pain's having intrinsic value? And analogous concerns attach to Bradford's view: if pains are normally intrinsically bad, *how* can enjoying them defeat this intrinsic badness? Philosophers cannot simply stipulate that the values that exist in the world happen to match our commonsense moral intuitions, as this gets the order of explanation backwards. (This is where the "theft" occurs.) When we talk as philosophers, quantifying over "intrinsic value", we are engaged in metaphysical and ontological discourse and we cannot simply assert that the metaphysics gives us what we want without telling us how we get it. Our commonsense intuitions are not supposed to be intuitions about what our intuitions are.

Evidently, our folk intuitions about pain and pleasure contradict each other. It cannot be the case both that pain is intrinsically bad and that the Type 2 Fun of alpinists is not bad at all (however unpleasant it may be), nor can it be the case both that pleasures are intrinsically good and that there is nothing intrinsically good about the assassin's pleasure. To ameliorate these contradictions, philosophers have continuously argued by *modus ponens*, adding iterated epicycles of explanation to their theories of either intrinsic value and/or pain and pleasure. In our view, a theoretically more powerful and elegant explanation arises if we argue by *modus tollens* and we reject the presumption that pains and pleasures have *intrinsic* value. That is the claim which is causing all the theoretical trouble. Rather, we ought to see the value of pain and pleasure as contingent, not necessary, and as instrumental and never intrinsic.<sup>12</sup> The inference of "it is bad" from "it is painful" is invalid. Pains can be intensely bad but are never intrinsically bad, and the same goes for pleasure and goodness. As laid out in the next section, it becomes easy to see why the assassin's pleasure is bad (and only bad) and makes what the assassin does only worse, and why the appropriate grief over the death of a loved one is only good, despite its being undeniably painful.

One lacuna requires attention. We have been assuming that the values of pain and pleasure are supposed to be moral values, as this is how various moral theories, such as utilitarianism have employed it. But those who want to hold onto the intrinsic values of pain and pleasure might try to claim that these values are actually prudential, and they might sit alongside whatever moral values are present in a situation.

<sup>12</sup> This might be read as a concession to the conditional views of intrinsic value, however, it is not: on conditional views, intrinsic value is attributed to the existence of certain conditions, rather than the intrinsic properties of the item or experience at hand. Rather, we hold that the value of pain/pleasure is contingent, depending on the context in which that pain/pleasure is felt. Thus, we do not try to explain pain's intrinsic badness by conditional means, but rather reject pain's intrinsic badness in favor of a contingent value that takes context into account.

So, adapting an example from Amartya Sen (1984, p. 209–10): imagine a physical ailment for which there are two treatments A and B. A has fewer side effects than B, so it has more intrinsic, prudential value for the patient, but A was tested on animals while B was not, so A has less intrinsic, moral value. This can be extended, of course, with regard to the value of Type 2 Fun, we might say that however intrinsically bad the pain of ice-climbing is, it is nevertheless intrinsically, prudentially valuable from the prudential point of view as it is part of what makes ice-climbers happy. And if this is the case, then perhaps all the examples above of valuable pains and disvaluable pleasures can be explained away in the manner just described.

In response, we might wonder how something intrinsically bad in any sense could genuinely be good for a person, such that the intrinsic badness of the pain of ice-climbing might nevertheless be part of the prudential happiness of the ice-climber. The first worry is that these are the sorts of situations which are typically thought of, when they obtain, as examples of masochism: the masochist enjoys the pain and is made happy by it or by what is intrinsically unpleasant.<sup>13</sup> If we argue by *modus ponens*, we should say this about ice-climbers. Perhaps this is plausible for ice-climbing but notice that the suggestion generalizes to the point of dissolving Rozin et al.'s (2013) category of benign masochism, such that all masochism is pathological: if the burn of the chili is in any way intrinsically bad, then lovers of spicy food are all pathological masochists for enjoying it. Moreover, the issue spreads beyond examples of benign masochism since, on such a view, enjoying any vexing challenge, like a difficult crossword puzzle, would be pathologically to some proportional degree, masochistic. The deeper worry, however, is analogous to Bradford's zero value objection: if the ice-climber's pain is intrinsically, prudentially good for them, then the more of that intrinsic value there is, the better the moral value of the situation should be. But this is not what we typically think: how can things which are intrinsically bad (in any sense) improve the moral situation?

The deeper, more substantial worry, however, is analogous to Bradford's zero value objection: we assume that instances of intrinsic goodness and badness are antagonistic. (How could it be otherwise?) If the pain of ice-climbers, however intrinsically bad it may be, is nevertheless prudentially, intrinsically good for them, then the more pain there is, the happier they should be. But this is backwards, since it implies that something which is intrinsically bad is enhancing or improving something which is intrinsically good. The worry becomes heightened when thinking about bad pleasures instead of good pains: consider the prudential value of the pleasure of the assassin and how it makes the assassination worse, not better. Or, worst of all: consider again the case of the sexual assault victim who experiences pleasure during the attack. On the view under consideration, the rape victim's situation should be improved by the intrinsic, prudential value good of the pleasure experienced during the attack, yet we've argued that this is the wrong read of the case.<sup>14</sup>

<sup>13</sup> We use "masochist" here in the colloquial sense not associated with a paraphilic disorder.

<sup>14</sup> Sen's example could also be read as saying that a single pain could have both instrumental value or disvalue and intrinsic disvalue or value. To attend to one option, consider bad pleasures, like that of the assassin, which would be intrinsically good for the assassin as it makes her happy, but it would nevertheless contribute to making the situation worse. Again: intrinsically good things are not supposed to make a

## 4 Explaining the Error and Its Prevalence

We offer the foregoing as an argument for thinking it is an error to see pain and pleasure as intrinsically valuable. Still, the questions of why the error occurred and what sort of value pain and pleasure may have remain. To resolve the error, anything written prior to 1859 is sufficiently outdated, as there is no denying that only since then has it been possible to see pain and pleasure as the result of evolution by natural selection (Darwin 1988). We now know that pain and pleasure evolved to serve biological—not axiological—functions, and we know that for pain this typically works in us, and in a wide variety of phyla, through nociception, which evolved at least 500 million years ago (Walters and Williams 2019). Pain and pleasure are reliable yet far from infallible feedback mechanisms signaling to a creature various states of its body in its present situation. Their purpose is to detect or diagnose situations, and guide or motivate responses (Melzack and Casey 1968). Animals feel physical pain as the result of exogenous harm to the body or endogenous malfunction.

Pleasure is originally the body's signal that something good for the organism is occurring and pain is the body's signal that physical harm is occurring. The fruit tastes sweet and is nourishing, the pleasure moves us to eat it; the fire is hot and harms flesh, pain is the detection of the harm and the cause of our withdrawal from the fire. Pain is often thought of as an "alarm system" (Dennett 1994, p. 61), so the "signal" of pain is like a warning light flashing or meter's needle hitting the "red line". As Barbara Finlay (2019) suggests in her comparison of visual and nociception systems, pain guides adaptive behavior which involves complex assessments of contingent relations between harmful stimuli and behavior.

The key difference between how the folk think of pain and how scientists do today is captured by Randolph Nesse and Jay Shulkin (2019) when they write, "Pain always seems like a problem, but it is usually part of the solution" (p. 1). It is reasonable to think pain itself is not a good thing, but it normally causes animals to move away from what is harmful which is good for them. Nesse and Shulkin make explicit a distinction missed by the folk: pain itself and pleasure itself are *not* subjects of natural selection but the capacities to feel pain and pleasure are. Natural selection has no stake in these sensations per se, we only evolved to feel them because of how they make us react. It is reasonable to conclude that the proper function of pleasure is to lead us to do what is good or salutary for the body and the mind, and the proper function of pain is to help us avoid what is bad or harmful. As such, it seems to us that Murat Aydede's (2014) functionalist theory of pain is the most promising extant view of pain.

---

Footnote 14 (continued)

situation worse; that is the point of the generic Platonic/Kantian conception of *intrinsic value* with which we begin Section III.

We should note that Sen's example was constructed to argue against there being an "all things considered" point of view that reconciles the differences between moral and prudential value. Arguing directly against this claim would take us too far from our topic, and it is not a claim that is typically embraced by those who claim that pain and pleasure are intrinsically valuable. Bloomfield (2014, 2017) has argued against the distinction between moral and prudential values.



This evolutionary genealogy serves as the starting point from which to see the axiological error we endeavor to expose. Again, leaving pleasure aside, and from the point of view of natural selection, pain has purely instrumental value.<sup>15</sup> We find one argument against this claim, made by Bradford, where she writes:

While [thinking that pain is only instrumentally valuable] may be plausible considering one's own pain, it is less plausible when we vividly reflect on the serious, excruciating and undeserved suffering of others, especially those whom we love. Even when excruciating pain is in the service of some further good and will eventually go away, it is clear that the pain is in a significant sense bad for the person who experiences it. This is such a widely held and plausible thought that if an account is in the business of capturing intrinsic bads, it would be a serious omission if it did not include pain (2021, p. 589).

We find this quote captures just the error we hope to have exposed because Bradford assumes that what is unpleasant for a person is bad for a person: what is clear is that the painful experience is unpleasant. We claim it is a further (invalid) inference to its badness. We need not deny pain's unpleasantness to deny its intrinsic value. We think that whether or not an unpleasant experience is good or bad is not determined by how it feels but by what it does. If the pain itself harms, as some intense pains can harm and traumatize, it is bad. But if the pain is part of a process of withdrawing from harmful stimuli or part of strengthening or healing, the pain is not bad, just unpleasant. A parent witnesses their child get a measles vaccine which is unpleasant for the child, but to call this experience "bad" for the child is to mischaracterize it. If our loved one is addicted to opioids and the withdrawal from this will be excruciating, we sympathize with our loved one as they experience this but do not think that it is bad for them, however unpleasant. The idea is not to deny the unpleasantness of the pain of the vaccine or the withdrawal, nor would we deny that painless treatments in general are preferable to the painful versions of them; we only deny that the pain itself is bad.

Pain normally serves to alert us to injury or harm so that we may address it; however, as a result of natural selection, it does so imperfectly. And so, we should expect "false alarms" or false positives. Some pains, like the pain of opioid withdrawal are the result of a body's recuperation. Other pains result from the body becoming healthier or stronger: an example of this kind of false positive is felt when someone who is not limber bends over with straight knees to touch their toes. There is no harm or injury in mild stretching, it is only healthy for muscles to be moderately stretched, and yet it is nevertheless uncomfortable if not painful. Like rabbits who are startled at the drop of a hat, from an evolutionary point of view, it makes sense to have an "injury-detection" system which is over-sensitive rather than

<sup>15</sup> And insofar as this pain functions to alert us to injury or illness, this is instrumentally good. While this is logically compatible with views of the intrinsic badness of pain, we have already argued against the view that pain has intrinsic disvalue. We suggest that this evolutionary story should be the starting point for axiology: the value of pain and pleasure are instrumental and dependent on their function. Any claim about their intrinsic value is a further theoretical posit, and we think, a pernicious one which has led us normatively astray. More on this in the conclusion.

under-sensitive, as the reverse could get us injured or killed. Unfamiliarity is often unpleasant, first registering as discomfort, if not pain. So, we should expect pain signals to be fallible, erring on the side of yielding false positives. (Of course, there are false negatives too, for there are a variety of maladies, e.g., cancer, which can go undetected for a long time.)

It is not mysterious why commonsense should conflate what causes pain with what is bad, and thereby conflate pain itself with badness itself, as commonsense is of course not in the business of making subtle axiological distinctions. As noted, from an evolutionary standpoint, we are “built” to be motivated by pain: this is part of pain’s proper function. It is not difficult to see how people might naturally infer that what we are built to avoid is, as it were, badness itself. And from there, the philosophical move to pain’s intrinsic badness can seem tautological, as witnessed in the quote from Goldstein (1989) above. It is not from commonsense, but only from an evolutionarily-informed point of view, that we can hermeneutically conceive, again, that “[p]ain always seems like a problem, but it is usually part of the solution” (Nesse and Shulkin 2019). Of course, philosophers started with commonsense, millennia before they learned about evolution by natural selection. So, the philosophical mistake is ancient and originally quite reasonable: pain reliably indicates the presence of something which is bad, and so it has been seen as badness itself.

Today, however, we know that pain evolved for its instrumental value, and so our axiology should keep up with our empirical knowledge. One might then reasonably ask, how could a trait – the ability to detect harm – with purely instrumental value yield a signal which itself has intrinsic disvalue? The answer in the literature from the Moorean point of view is that the badness of pain is non-natural and unanalyzable, which simply begs precisely the question at hand. A more sophisticated answer comes through the conditional theory of intrinsic value. In defending such a theory, Shelly Kagan (1998) suggests that, in at least some cases, “the intrinsic value of an object depends completely upon its instrumental value” (p. 286). In a footnote on the same page, he walks back the “completely”, acknowledging that the instrumental value may be only necessary but not sufficient for the intrinsic value. Nevertheless, the thesis of his essay is that relational properties can be relevant to something’s intrinsic value. Unfortunately, we do not find Kagan’s examples of the intrinsic value of luxury goods persuasive (i.e., sports cars for Kagan and mink coats for Korsgaard (1983)), nor is Kagan’s idea that something’s uniqueness (a relational property) can be the source of intrinsic value. (His example of the intrinsic value of the pen used by Lincoln to sign the Emancipation Proclamation seems to conflate intrinsic value with projected sentimental value: that pen is valuable only in the way that a child values a teddy bear and a patriot values a flag.)

What is relevant for our purposes is to ask why we should think that pain’s instrumental value constitutes or partly constitutes pain’s putative intrinsic value. And the argument would presumably have to be because we avoid pain (and seek pleasure) for its own sake, as an end in itself (cf. the discussion of Korsgaard (1983) above on seeking pleasure for its own sake.) But, once again, this begs exactly the question we are asking, as our thesis is that, however widespread, it is a moral mistake to avoid pain or seek pleasure for their own sake specifically because they have no intrinsic value. (More on this below.) Despite attempts like Kagan’s (1998) and

Korsgaard's (1983), it is far from obvious how to derive intrinsic value from a fallible biological mechanism, the *raison d'être* of which is to instrumentally signal the (putative) detection of what is harmful (or salutary) in the world; recall from Nisse and Shulkin (2019) above that the capacities to feel pain and pleasure are the results of natural selection but that pain and pleasure themselves are not. This is not to say that an axiological conversion from instrumental to intrinsic value is logically or conceptually impossible, but it would need substantial defense and cannot simply be assumed or taken for granted.<sup>16</sup> At the very least, the intrinsic disvalue of pain is not analytic nor true a priori.

This seems all the more true in the face of the counterexamples discussed above to the idea that pain is intrinsically bad. Rozin et al. (2013) explain benign masochism as a kind of "hedonic reversal" by which a usually negative experience is converted into a positive experience. (An example of the other direction is the way that being tickled can quickly go from fun to torment.) They write:

Benign masochism refers to enjoying initially negative experiences that the body (brain) falsely interprets as threatening. This realization that the body has been fooled, and that there is no real danger, leads to pleasure derived from "mind over body." This can also be framed as a type of mastery (p. 439).

We return below to the point about mastery. For now, notice the error alluded to by Rozin et al.: the body's alarm system is giving a false positive. The experiences causing the pain are not in fact harmful and so, with habituation, we can begin to enjoy them. Yet if, according to commonsense, these painful but non-harmful experiences are intrinsically bad, it is hard to conceive how they could possibly reverse their "polarity", so to speak, and become intrinsically good over time. However tentitious it may be to suggest, as defenders of the conditional theory of value do, that something without intrinsic value could gain it by becoming a part of something intrinsically valuable, it is quite another to suggest that something intrinsically bad could become intrinsically good just because we have habituated to it, as is the case with hedonic reversals.

Famously, the issues here are vexed. Daniel Dennett's argument in "Quining Qualia" (1988) depends on these sorts of hedonic reversals. Recall Dennett's story about Chase & Sanborn and the flavor of coffee: no one likes coffee at first, it tastes bitter and bad on the first sip, but after repeated exposure, it tastes good. So, Dennett asks, does the flavor itself change from bad to good or does the flavor stay the same and we merely perceive that flavor first as bad and then as good? While Dennett's conclusion is that we should dispense altogether with the qualia, the point here does not involve the ontology of qualia or mental states in general. Rather, our point is axiological: we note how difficult it would be to explain how something that is

<sup>16</sup> Cf. Plato on justice and borrowing and Kant on good will. The way we see intrinsic value is consistent with an item's possessing it in virtue of its functions or how it works, but what would make an item intrinsically, and not merely instrumentally, valuable is that it never fails regardless of circumstance or context. If an item functions well in all possible circumstances, it seems reasonable to think its value is intrinsic and not merely instrumental.

intrinsically bad can become intrinsically good just by being repeated a number of times. If pain and pleasure are (respectively) intrinsically bad and good, then are we supposed to imagine that the taste of the coffee or the burn of the chili starts off being intrinsically bad and then, somehow, metamorphoses into something intrinsically good? Even if this axiological mystery could be solved, arguing by *modus ponens* will have taken us a long way from the sort of supposedly “uncontroversial claims” with which we began.

Importantly, these kinds of hedonic reversals occur in cases that have nothing to do with either taste or benign masochism (at least as defined by Rozin et al.). For someone who is unwilling, one of the most unpleasant things to do is exercise. It is uncomfortable, it feels just horrible. And yet, that very same activity, if sufficiently pursued and repeated, can become so pleasant as to be something to which a person can become “addicted”, at least in the sense of that word which is apt when people feel a need to exercise and feel badly when they do not do it.

Let us go back to our stiff and unlimber person who is trying to bend over to touch their toes with straightened knees. The person would experience discomfort or even twinges of pain, in the hamstring and lumbar muscles, and in the back of the neck. All yogis recognize that strength and flexibility will not develop without experiencing what is unfamiliar and uncomfortable, and some distinguish “good pain” from “bad pain”. Discomfort or good pain builds slowly during exercise and stops when the exercise is finished, while bad pain, which is the result of injury, begins suddenly during exercise and continues after the exercise is over or it begins immediately after the exercise is finished. Whether or not a pain is good or bad depends on whether they are the result of a salutary or harmful process. Evidence for this good pain/bad pain distinction is that movements or exercises which cause discomfort or good pain cease to be painful if repeated, as strength and flexibility are developed, while this is obviously not true for bad pain. Strengthening weakness is always uncomfortable and sometimes painful but nevertheless instrumentally good for a person, so it would behoove us to be able to differentiate this kind of pain from bad, harmful pain, first theoretically then phenomenologically and, finally, to master good pain as an “acquired taste”, as the vexations of difficult crossword puzzles are an acquired taste. Mastery of this kind of good pain leads to steady improvement.<sup>17</sup>

And this brings us back to Rozin et al.’s point about “mastery” (2013, p. 439; quoted four paragraphs up). When we engage in activities which are new to us, they feel unfamiliar, thinking again of the unlimber person stretching, and our bodies react to unfamiliarity as if the motion is dangerous, activating the “alarm system” of pain. But as we become familiar with the activity, the body learns that it is actually safe, and the activity can become enjoyable. We become “masters” of stretching muscles, spicy chilies, roller coasters, etc., and as we do so, we begin to enjoy and feel pleasure at what was originally quite the opposite. This kind of “hedonic conversion” is possible only if the value of pain and pleasure are *not* intrinsic.

<sup>17</sup> Thanks to Sheri Trellevik for discussion on the topics in this paragraph, while acknowledging her thought that the term “good pain” may not be helpful for students of yoga.

Joining the evolutionary story of pain and pleasure with these counterexamples to pain's intrinsic value, we return to the simple way the original and commonsense axiological error arose in the folk and yielded the false conclusion that pain is intrinsically bad and pleasure is intrinsically good. Pain is the body's general signal to the mind that something bad is happening and pleasure is the general signal that something good is happening. We treat the signal of something bad as being bad in itself; but this is like shooting the messenger for bearing bad news. And intense pains, traumatizing pains, are certainly bad. As noted by Neese and Shulkin (2019) "[p]ain always seems like a problem...". The error comes from inferring from what is commonly and generally true to conclusions about what is necessarily (or even analytically) true. However bad some pains may be, all pains are not necessarily bad. Such an inference is invalid, and this is why counterexamples arise and our intuitions contradict one another.

This simple story is the reason for the philosophical dogma and the explanation of the error behind it: prior to having adequate empirical knowledge of the phenomenon, prior to Darwin, commonsense mistook a generally true relation for a necessarily true one and this error was philosophically petrified by conflating something that is only instrumentally valuable with something that is intrinsically valuable. The error is exposed when we recognize that, for instance, exercise is painful for someone out of shape even though it is salutary and good for that person; smoking cigarettes is in fact bad for a person even though it "feels good" and is pleasurable. However pleasant it may be, there is nothing good about the assassin's pleasure, and however unpleasant it may be, there is nothing bad about grieving appropriately. These examples show how some pains can actually be good and how some pleasures can be bad, and thus why pleasure and pain are not intrinsically valuable. While pain is always, necessarily, analytically unpleasant, it is actually (instrumentally) good when it gets us to avoid what is bad for us or when it is part of a healing or strengthening process. Pain is only (instrumentally) bad when it keeps us from doing what is good for us or when it is a false positive signal of harm. While pleasure is analytically pleasant, it is only (instrumentally) good when it motivates us to do what is good for us and it is (instrumentally) bad when it leads us to do what is bad for us.

## 5 Conclusion

The moral errors that have been a result of this axiological error about pain and pleasure are numerous. Consider how broad the scope of moral choice is which involves choosing between two good options, or choosing the lesser of two evils, or any consequentialist trade-offs between what is intrinsically good and bad. Now, consider how many of these choices are made while thinking that pleasures automatically count as good and pains automatically count as bad. If we started denying the goodness of pleasure per se and the badness of pain per se, if we started thinking of these merely as (un)pleasant but without intrinsic value, the effects on our choices would be enormous. If we taught ourselves the proper functions of pain, if we cease to be fooled by the supposed "intrinsic disvalue" of it, so that we do not automatically think that experiencing pain is experiencing something bad, then we would not

be fooled into avoiding pains which we know are not harmful by telling ourselves that the pain itself justifies us in avoiding it.

This may be true at the level of common sense, but the changes to moral theory and ethical choice become stark. Typically, when we do something we know is a little bit wrong because of a large amount of pleasure we get from doing it, a great deal of moral theory tells us that the (putative) good of the pleasure can outweigh the badness of the act, so the overall right thing to do is the bad thing. If, in such a case, we take away any thought that the pleasure, however pleasant, is itself intrinsically good, then our deliberations ought to change accordingly: we are morally and ethically not bound to do the small wrong thing. If this is true for small bad acts, it is true *a fortiori* for worse acts than that: though there may be exceptions to the rule, pleasure itself does not normally justify doing something morally wrong. Pleasure can be a reason to do what is good, this is why we can experience pleasure in the first place, but it cannot be a reason to do what is bad. *Mutatis mutandis*, the same holds true for pain: discomfort, or the unpleasantness of pain, is never, in itself, an excuse to avoid doing something morally good.

Consider what the world would be like if the present view of pain and pleasure were fully accepted, and our upbringing and behavior were trained to act in accord with it. This would mean that pain and pleasure would no longer be seen as intrinsically valuable, and that the differences between good pains and bad pains, and good pleasures and bad pleasures would be learned and inculcated into our lives. We would learn to avoid pain when it causes harm but learn not to avoid it when we know this will lead to strength and flourishing; we would learn to pursue pleasure when it is salubrious and leads to good health and to avoid pleasure when it is harmful or enervating.

The claim here is not that this axiological knowledge would rid us of immoral behavior. Pleasure will remain as pleasant and tempting as it always has been; pain will never lose its unpleasantness. But changing the way we think about pleasure and pain will change our moral deliberations and choices. Weakness of will or akrasia will not cease to exist, but akratic people will no longer be able to fool themselves into thinking that there is something intrinsically bad about doing something unpleasant or something intrinsically good about experiencing a bad pleasure.

Unsurprisingly, the greatest change would be to hedonic theories based on the goodness of pleasure and the badness of pain. A form of utilitarianism would remain possible, but the *summum bonum* to be maximized would not be understood directly in terms of pleasure and pain *simpliciter*. Rather, it would involve the maximization of good pleasures and good pains and the minimization of bad pleasures and bad pains. Given the present view, all traditional forms of hedonism are founded on ancient, empirical mistakes about the nature of pain and pleasure which were understandably ubiquitous before Darwin. There is, in fact, *no* moral reason to pursue something bad yet pleasurable or to avoid something good yet unpleasant. We may still make the errors in moral thought but we may at least avoid this basic error, and we may thereby become better at pursuing what is truly good and avoiding what is truly bad.

Those are the practical effects of the present view. From a theoretical perspective, accepting the view that pain and pleasure have only instrumental value is far more

theoretically simple, powerful, and elegant than all the epicycles of theory required to preserve the ancient yet mistaken belief that pain is intrinsically bad and pleasure is intrinsically good.<sup>18</sup>

**Competing interests** The authors declare none.

## References

- Annas, J. 1993. *The Morality of Happiness*. New York: Oxford University Press.
- Aydede, M. 2014. How to Unify Theories of Sensory Pleasure: An Adverbialist Proposal. *Review of Philosophy and Psychology* 5: 119–133.
- Bentham, J. 1789. *An Introduction to the Principles of Morals and Legislation*. London: T. Payne and Son.
- Bloomfield, P. 2017. Morality is Necessary for Happiness. *Philosophical Studies* 174 (10): 2613–2628.
- Bloomfield, P. 2014. *The Virtues of Happiness*. New York: Oxford University Press.
- Bradford, G. 2020. The Badness of Pain. *Utilitas* 32: 236–252.
- Bradford, G. 2021. Perfectionist Bads. *The Philosophical Quarterly* 71 (3): 586–604.
- Bradley, B. 2002. Is Intrinsic Value Conditional? *Philosophical Studies* 107: 23–44.
- Darwin, C. 1888. *On the Origin of the Species*. New York: New York University Press.
- Dennett, D. 1994. *Consciousness Explained*. New York: Little, Brown.
- Dennett, D. 1988. Quining Qualia. In *Consciousness in Contemporary Science*, ed. A. Marcel and E. Bisiach. New York: Oxford University Press.
- Epicurus 1994. *The Epicurus Reader*, ed B. Inwood and L. Gerson. Indianapolis: Hackett Publishing Company.
- Finlay, B. 2019. The Neuroscience of Vision and Pain. *Philosophical Transactions of the Royal Society B* 374: 20190292. <https://doi.org/10.1098/rstb.2019.0292>.
- Finlay, S. 2014. *A Confusion of Tongues*. New York: Oxford University Press.
- Feldman, F. 2004. *Pleasure and the Good Life*. Oxford: Clarendon Press.
- Frankena, W. 1973. *Ethics*, 2nd ed. Englewood Cliffs: Prentice Hall.
- Goldstein, I. 1989. Pleasure and Pain: Unconditional, Intrinsic Values. *Philosophy and Phenomenological Research* 50 (2): 255–276.
- Hurka, T. 1998. Two Kinds of Organic Unity. *The Journal of Ethics* 2 (4): 299–320.
- Kagan, S. 1998. Rethinking Intrinsic Value. *The Journal of Ethics* 2 (4): 277–297.
- Kant, I. 1997. *Groundwork of the Metaphysics of Morals*. M. Gregor, Translator. Cambridge: Cambridge University Press.
- Korsgaard, C. 1983. Two Distinctions in Goodness. *The Philosophical Review* 92 (2): 169–195.
- Kraut, R. 2007. *What is Good and Why: The Ethics of Well-Being*. Cambridge, MA: Harvard University Press.
- Lemos, N. 1994. *Intrinsic Value: Concept and Warrant*. Cambridge: Cambridge University Press.
- Mayerfeld, J. 1999. *Suffering and Moral Responsibility*. New York, NY: Oxford University Press.
- Melzack, R., and K. Casey. 1968. Sensory, Motivational, and Central Control Determinants of Pain. In *The Skin Senses*, ed. D. Kenshalo. Springfield, IL: Charles C. Thomas Publisher.
- Mill, J.S. 2001. *Utilitarianism and the 1868 Speech on Capital Punishment*, ed. G. Sher, 2nd ed. Indianapolis, IN: Hackett Publishing Company.
- Nakamura, J., and M. Csikszentmihalyi. 2009. The Concept of Flow. In *Oxford Handbook of Positive Psychology*, ed. C. Snyder and S. Lopez. Oxford: Oxford University Press.
- Nesse, R., and J. Schulkin. 2019. An Evolutionary Medicine Perspective on Pain and its Disorders. *Philosophical Transactions of the Royal Society B* 374: 20190288. <https://doi.org/10.1098/rstb.2019.0288>.

---

<sup>18</sup> We are grateful for comments on earlier drafts of the essay from Gwen Bradford, Ben Bradley, and Philip Ebert, and comments from an anonymous referee.

- Norcross, A. 2004. Puppies, Pigs, and People: Eating Meat and Marginal cases. *Philosophical Perspectives* 18 (1): 229–245.
- Plato. 1993. *Republic*. Translated by Robin Waterfield. Oxford: Oxford University Press.
- Plato. 1992. *Protagoras*. Translated by S. Lombardo and K. Bell. Indianapolis: Hackett Publishing.
- Rozin, P., L. Guillot, K. Fincher, A. Rozin, and E. Tsukayama. 2013. Glad to be Sad, and Other Examples of Benign Masochism. *Judgement and Decision Making* 8 (4): 439–447.
- Samet, M. 2011. *Climbing Dictionary*. Seattle, WA: The Mountaineers Books.
- Sen, A. 1984. Well-Being, Agency and Freedom. *The Journal of Philosophy* 82 (4): 169–221.
- Sidgwick, H. 1981. *The Methods of Ethics*, 5th edn. Indianapolis IN: Hackett Publishing Company.
- Singer, P. 1975. *Animal Liberation: A New Ethics for Our Treatment of Animals*. New York, NY: The New York Review/Random House.
- Solomon, R. 2004. *In Defense of Sentimentality*. Oxford, UK: Oxford University Press.
- Sumner, W. 2020. The Worst Things in Life. *Grazer Philosophische Studien* 97: 419–432.
- Tiberius, V. 2018. *Well-Being as Value Fulfillment: How we can help each other to live well*. Oxford: Oxford University Press.
- Walters, E., and A. Williams. 2019. Evolution of Mechanisms and Behavior Important for Pain. *Philosophical Transactions of the Royal Society B* 374: 20190275. <https://doi.org/10.1098/rstb.2019.0275>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.