The Agony of Reason: The Unsteady Bond Between Suffering and Human Rationality

Matthew Fulkerson† and Jonathan Cohen‡

Other creatures live in the world more or less as Nature presents it to them; and they react to it more or less directly, albeit sometimes with remarkable sophistication. In contrast, we human beings live to a significant degree in the worlds that our brains remake — though brute reality too often intrudes.

— Ian Tattersall, Masters of the Planet: the search for our human origins

On the face of things, it would seem that suffering can have significant value for creatures like ourselves. It is overwhelmingly plausible from the point of view of both naïve reflection and broadly empirical investigation that suffering is a necessary condition on a certain kind of fitness-preserving causal motivation, and one that assists and supports rational activity. On such a view, suffering is a primary source of “motivating reasons,” viz., practical reasons that guide a subject’s actions.1 Thus, when Lucy the distracted machinist strikes her thumb.

1Practical reasons are often contrasted with “justifying reasons,” those considerations that justify or give good reason for an action, whether the agent acts on those considerations or not. While we believe that suffering provides both motivating and justifying reasons for action, in what follows we will focus only on the motivating reasons provided by suffering.

That said, we do not wish to commit to any particular account of action explanation or motivating reason in what follows, since we believe the interest and importance of the disconnect we highlight below is independent of such controversies. Thus, we will sometimes characterize motivating reasons in terms of instrumental belief-desire pairs, and other times we’ll talk about states of suffering themselves as motivating reasons, as best smooths overall exposition. Readers with more committed views are invited to substitute their preferred ways of describing reasons as appropriate. (If, for example, one prefers not to talk about a state of suffering itself as a motivating reason, one may take our use of such states/facts as a shorthand to pick out the genuine motivating reason — e.g., to pick out

---

*This work is fully collaborative; authors are listed anti-alphabetically.
†Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, mfulkerson@ucsd.edu
‡Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, joncohen@aardvark.ucsd.edu

1Practical reasons are often contrasted with “justifying reasons,” those considerations that justify or give good reason for an action, whether the agent acts on those considerations or not. While we believe that suffering provides both motivating and justifying reasons for action, in what follows we will focus only on the motivating reasons provided by suffering.

That said, we do not wish to commit to any particular account of action explanation or motivating reason in what follows, since we believe the interest and importance of the disconnect we highlight below is independent of such controversies. Thus, we will sometimes characterize motivating reasons in terms of instrumental belief-desire pairs, and other times we’ll talk about states of suffering themselves as motivating reasons, as best smooths overall exposition. Readers with more committed views are invited to substitute their preferred ways of describing reasons as appropriate. (If, for example, one prefers not to talk about a state of suffering itself as a motivating reason, one may take our use of such states/facts as a shorthand to pick out the genuine motivating reason — e.g., to pick out
hard with a ball-peen hammer, or when Luke the graduate student is compelled to overcome his shyness to lead a seminar discussion, it is extremely natural to describe the situations by saying that their forms of suffering provide Lucy and Luke with reasons to act in certain ways. (For Lucy perhaps pain gives a reason to tend to her injured thumb; for Luke perhaps social embarrassment gives a reason to focus his attention on the intellectual material at hand.) These, like similarly ordinary descriptions of everyday cases motivate treating the picture of suffering as reason-conferring as a default starting place.

This picture gathers support from the observation that suffering and reasons hang together in surprisingly resistant ways, even in cases where you might have expected the two to come apart. Thus, for example, it is interesting that pain asymbolics and those with severe leprosy tend to ignore proposed replacement harm signals (say, ringing bell sounds or flashing lights) unless the intensity of these signals is increased to a point that induces suffering: it would appear that signals that fail to induce suffering just do not compel motivationally/rationally in the way that suffering does (Auvray et al. 2010; Brand and Yancey 1993). Or, again, it is interesting that canonical descriptions of learned helplessness (e.g., in clinical depression or as a result of uncontrolled stress) involve both a lack of affect and a lack of motivation (Abramson et al. 1978). In a similar vein, it is interesting that cornered prey animals both freeze (stop being compelled to act) and release pain-masking endogenous opioids (thereby presumably mitigating suffering) just at the time when suffering ceases to be adaptively or motivationally useful (Amit and Galina 1986).

As we say, we are inclined to accept such pretheoretical and broadly empirical considerations at face value, and to accept the picture of suffering as reason-conferring that they appear to support. However, accepting this picture invites the set of motivating attitudes typically caused by that state (cf. Smith 1994). In the case of pains, these would be those beliefs and desires intrinsic to typical painful experiences.)

Interestingly, the suffering caused by the alarms do not motivate subjects to avoid the genuine harms indicated by the alarms; instead, they strongly motivate behavior to shut off or remove the alarm signal.

Similar effects have been found in humans; see e.g., the discussion of the human pain inhibitory system in Hardcastle (1999, 132–134).

In saying that the picture of suffering as reason-conferring enjoys support both from our pretheoretical picture of the world and broadly empirical investigation, we do not mean to say that there are no puzzles or worries that can be raised against it. For example, there is an expanded version of the worry about heterogeneity that has been raised in connection with pain (Heathwood 2006; Smuts 2010): if concerns about heterogeneity motivate skepticism about the prospects for systematic theorizing about pain, those skeptical worries are presumably yet more pressing as applied to the far more inclusive category of suffering. Or, again, if you are inclined to think about pain and/or certain forms of suffering in causal terms, sympathy to certain views about the relation between
a concern that, to our knowledge, hasn’t been discussed elsewhere, over the nature of and limits to the interaction between suffering and rationality.

One way to bring out the disconnect we have in mind starts with the observation that states of suffering play a positive rationalizing role in our actions — they are *reason conferring* — although they seem immune or resistant to rational considerations — they are not *reason-responsive*.

Now, suffering is plausibly not unique in this respect. Many states, events, facts, and states of affairs can plausibly serve as reasons for a subject without themselves being responsive to reason. For instance, if Theo is thirsty and heads to the fridge because he knows there is water there, there’s a good, if minimal, sense in which the water’s being in the fridge counts as a reason for his action; but of course, the water’s being in the fridge is completely immune to influence by Theo’s broader reasons and motivational profile.

What makes the case of suffering plausibly different from and more interesting than the case of the water’s being in the fridge (etc.), however, is that the rational role of suffering for Lucy and Luke is significantly more extensive than that of water’s being in the fridge for Theo. The claim that suffering confers (practical) reason can be understood as the claim that suffering plays a central and often beneficial role in practical deliberation for the purposes of action (broadly construed). It is this point that makes our worry about the disconnect between suffering and reason so pressing.

Roughly, the worry is that, if suffering does confer practical reason, and thereby comes to be integrated (as we have seen, in central and important ways) with practical rationality in creatures like us, there seem nonetheless to be significant limits to that integration. In particular, we want to say that when suffering becomes (practical) reason-conferring for a subject (when it enters into what Sellars (1956) famously calls the “space of reasons”), it will nonetheless continue to behave in ways that distinguish it from other objects of rationality, and make it appear more like (yet nonetheless still distinct from) a kind of non-rational influence. In this sense, suffering seems to play a dual (partly rational, partly arational) role in the mental lives of creatures like us that it does not in the mental lives of psychologically less or more rationally sophisticated creatures.

We’ll begin by attempting to characterize this disconnect (§1), and arguing that it has serious negative psychological consequences (§2). Next we’ll ask how best to understand the disconnect. We’ll argue against several attempts to explain it away reductively in terms of more familiar rational psychological pathologies (§3), and put forward an alternative descriptive conception that treats the disconnect as a reflection of our peculiar, partly but not wholly integrated, rational psychology (§4). Finally, we’ll conclude (§5).

causes and reasons might make you doubt that pain or suffering could be a bona fide reason. (We address this latter worry in Cohen and Fulkerson (2013).)
1 A disconnect between suffering and action

To bring out the salient aspects of the disconnect between suffering and our rational psychologies, it will be useful to consider the contrast between the two following cases of practical deliberation, one of which involves suffering, and the other of which does not.

Beach: You are deciding whether to undertake the action of going to the beach for the afternoon. You desire an afternoon of relaxation, and you believe that going to the beach will achieve that end. This belief-desire pair, which does not involve suffering, and which we will call $R^+$, is a reason for you, and one that speaks in favor of the action of going to the beach. In addition, you have a looming deadline for submitting a paper, and you desire to fulfill this obligation. You believe that an afternoon spent working will achieve this end. This second belief-desire pair, which also does not involve suffering, and which we will call $R^-$, is also a reason for you, and one that speaks against undertaking the action of going to the beach.

In this situation you have two competing (non-suffering) reasons, $R^+$ and $R^-$, pointing in opposite directions with respect to the action; your decision how to act is an exercise of your rationality, pitting reason against reason. If one desire is stronger than the other, and no other considerations arise (it doesn’t start to rain or you don’t get an extension on your submission), the rational thing for you to do would be to satisfy that desire (again, on a rationalizing but not necessarily on a justifying account of action explanation). Assuming your desires are equally strong, and no other competing considerations arise, you might proceed by an explicit process of rational deliberation, comparing the relative utilities of the outcomes, looking for alternative ways of satisfying both desires (writing at the beach!), or reassessing your desires in light of the new evidence (it may rain, better to hold off on the beach day). At some point, this process of deliberation will end. In this case, you come to the considered view that the reason $R^+$ outweighs $R^-$, and so decide, decisively, by this exercise of rational deliberation, to go to the beach.

At that point, after having decided in favor of undertaking the action, what can we say about your view toward the reason $R^-$ that spoke against that action before you had undertaken it? Of course, $R^-$ may continue to be a pro tanto reason against acting. After all, you still have a deadline coming up, and a desire to fulfill your obligation to submit your paper. You also still have the beliefs about (at least) one possible course of action that would fulfill that desire. Now, if your decision was made simply because your desire in $R^-$ was not as strong as the one in $R^+$, then you are likely happy with your choice and will proceed directly to the beach without looking back. If the decision was more difficult, you might still feel some pull toward the afternoon of work, even as you head off to the beach. That said, it is notable that the exercise of your rationality by which you weighed $R^-$ against other reasons had the net effect of making $R^-$ cease to be an all things considered guide.
to action. After all, if you are on your way to the beach, you’re not acting on $R^-$. After deciding, you can consider and reflect consciously on $R^-$, and even recognize clearly its rational and motivational force, but it is no longer playing any active role in the explanation of your present actions. You have, by a successful exercise of rationality, converted $R^-$ from a reason that is potentially action-guiding to a reason that is not action-guiding.

**Vaccination:** A new vaccine has just hit the market. It is very effective at preventing a range of disorders, and thereby positively increases your long-term health. It requires that a subject remain completely still for two minutes while the drug is administered via an intravenous needle. If you move too much or remove the needle early the procedure must be repeated from the beginning. You are in the doctor’s office and, just before the shot is administered, are deciding whether to go through with the procedure. You have a longstanding desire to avoid serious illness and you believe (correctly) that getting the vaccination will help you achieve this end. You thus have one reason $R^+$ speaking in favor of the action of remaining still and going through with the procedure. On the other hand, getting a shot hurts, and you also have a longstanding desire to avoid pain whenever possible. You believe avoiding the shot will help you achieve this goal. You thus have a reason $R^-$ speaking against undertaking the action.

Now, since you are already in the doctor’s office, it is likely that your desire to have good health is stronger than your desire to avoid pain, and so the overall rational thing for you to do is to go through with the shot. But there is pain involved, so you’re having a much more difficult time sitting still while the nurse prepares the needle. As before, you embark on explicit deliberation, comparing likely outcomes, thinking of alternative considerations, or other ways to have it both ways. As before, you have two reasons, $R^+$ and $R^-$, pointing in opposite directions with respect to the action; your decision how to act is presumably an exercise of your rationality, pitting reason against reason. Suppose that you come to the considered view that, despite the pain, the reason $R^+$ outweighs $R^-$, and so decide, decisively, by this exercise of rationality to remain still and go through with the shot.

At this point, after having decided in favor of undertaking the action, what can we say about your view toward the reason $R^-$ that spoke against that action before you had undertaken it? From the point of view of rationality, it seems that this case is just like Beach. As before, $R^-$ plausibly continues to be a pro tanto reason against acting. As before, through the exercise of your rationality by which you weighed $R^-$ against other reasons you can make $R^-$ cease to be an all things considered guide to action. While you still have the desire to avoid pain, and you still believe moving away will help you achieve that goal, you have decided on a different course of action. As before, after having decided, you can consider and reflect consciously on $R^-$, and even recognize clearly its rational and motivational
force, without being tempted to change your mind about the decision, and without taking it as a guide for your action. You have, by a successful exercise of rationality, converted $R^\ast$ from a reason that is potentially action-guiding to a reason that is not action-guiding.

Thus, it would seem that the case of Vaccination and the case of Beach are, from the point of view of rationality, structurally analogous.

But, crucially, and unlike the case of Beach, where the reasons did not involve suffering, there is a way in which $R^\ast$ persists, or resists being overwhelmed by rational control. Namely, $R^\ast$ continues to exert control over our actions even after rational deliberation has done its work. Even if we manage to stay still, doing so requires continual executive control. We have to, as it were, “fight through” the pain in order to execute the action we have settled on — even after having settled on it.

---

5Michael Brady (personal communication) suggests that this characterization might not be true of at least some cases of emotional suffering, like suffering from guilt, shame, or regret, which he thinks depend on a cognitive evaluation and therefore can be subjected to rational control — e.g., when I am rationally convinced to give up the belief that I am responsible for the situation about which I feel guilt/shame/regret.

We have three points to make. First, if Brady’s cases of suffering are indeed subject to rational control, they are plausibly special — viz., unlike other instances of suffering — in just this way, which would mean that our story remains intact as a characterization of the general case. Second, Brady’s claim that states of emotional suffering are susceptible of rational control is controversial; one piece of evidence of this is that rational cognitive behavioral therapy and related interventions that bring rational considerations to bear appear to be effective with respect to a wide range of problems (especially when combined with medical and other forms of treatment), but have been found to produce more limited therapeutic effects on states of emotional suffering in the general population (cf. [Hofmann et al. (2012) 436]; [Davidson et al. (2006) 7]). Third, even if rational considerations can remove states of emotional suffering that depend on cognitive appraisal, it’s certainly possible to deny that, in such cases, the states of suffering themselves (as opposed to the cognitive states on which they depend) are under the influence of rationality. At a minimum, then, we deny that Brady’s cases impugn the story we are telling.

6A similar finding shows up in neuroscientific models of conflicts in decision-making between abstract long-term goals (say, a desire for health) and immediate short-term affective response (say, the pleasure of eating ice cream) (Teuscher and Mitchell [2011]; Bechara et al. [1994, 1999]; McClure et al. [2004]; Ballard and Knutson [2009]). As many of us know all too well, the immediate pleasure of ice cream ordinarily exerts a much stronger pull on action than the abstract long-term goal of health (MacKillop et al. [2011]). To be sure, we sometimes succeed in acting in accord with our abstract long-term goals. But when we do, this is not because our top-down long-term reasons gain in strength; rather, they appear to triumph by suppressing their short-term affective competitors (Benedetti et al. [2010, 2006]). This is consonant with our general picture in suggesting that top-down reason has a limited effect on suffering, and only by removing it. (To be fair, the same idea also cuts against our picture somewhat by suggesting a mechanism by which our disconnect can be resolved; but because these cases of successful top-down mediation appear to be
It would appear that the action-guiding force of the suffering-involving reason $R^\ast$ in Vaccination survives the exercise of rationality — and in this sense is interestingly disconnected from rationality — in a way that the force of the non-suffering-involving reason $R^-$ in Beach does not. Nor is this an artifact of this particular case; we all have experience of pain and suffering impacting our actions in ways that resist rational mitigation. In what follows, we want to argue that this disconnect holds important lessons about the nature and role of suffering in the mental lives of creatures like ourselves.

2 The agony of disconnect

One of the reasons that the disconnect between rationality and suffering we’ve identified bears interest is that, we claim, it has the potential to initiate novel forms of suffering (both sensory and psychological), and thereby can significantly impede the overall well-being of individuals subject to it. It also seems to reveal something about the particular nature and strength of our rational lives in an especially illuminating way that promises interesting theoretical implications. In this section, we’ll focus on the first claim; we’ll return to the second at the end of the paper.

Human suffering is often rooted in bodily reactions, but it is often mediated by the unique rational perspective we can take with respect to that suffering. Thus, even a brief examination of the literature on pain and other forms of human suffering suggests a strong role for top-down cognitive effects on our pain experiences. For instance, anticipation effects play a role in mediating elicited pain response; social context can influence pain judgements; and will power is inversely correlated with pain tolerance (Ossipov et al. 2010; Silvestrini and Rainville 2013). Limited, effortful, and temporary, we are inclined to treat them as exceptions that prove our more general rule.

An ordinary example of this kind involves spicy food. Capsaicin works by binding to TRPV1 receptors in the nose, mouth, and skin that are ordinarily activated only by high temperatures, thereby tricking these receptors into producing signals of heat at temperatures much lower than ordinarily necessary for their activation. But learning that the TRPV1 signals present after eating modestly spicy food are false positives — that there is no harm in the offing — does nothing to lessen the associated suffering.

A slightly more obscure example of this sort involves the “thermal grill” demonstrated by Thunberg (1896) — an alternating pattern of warm (40°C) and cool (20°C) bars that produces an experience of intense burning to the touch. Subjects know that the temperature range on the bars of the grill is well within the limits of safety, and can verify this by touching the warm and cool bars separately without undergoing any suffering. Yet the burning sensation produced by touching a larger region of the grill is so intense that subjects are unable to keep their hands on it (for a simple model of the mechanism underpinning our response to the thermal grill, see Craig and Bushnell 1994).
One area of intense investigation that most clearly reveals the potential for the kind of unique suffering we propose here are cases of chronic pain, in particular, chronic pain arising from conditions like central and peripheral neuropathy. Neuropathy refers in general to pains arising from nerve damage. Neuropathy can cause a wide range of painful sensations depending on the type of nerve affected and the nature of the damage. These can be intense and unpleasant tingles, shock sensations, burning, chills, and severe aches. The nerve damage can arise from many different causes, including diseases like diabetes, from direct trauma to the nerves, and can even be inherited. Neuropathy is remarkably and frustratingly difficult to treat effectively.

One of the most insidious aspects of chronic pains arising from neuropathy is their pervasive psychological impact. In many cases, the nerve damage itself raises few additional health concerns. Subjects who suffer this sort of pain know that it is “just” pain, and that it does not pose (in itself) any additional health risks. Such pains do not serve any protective or informative end, and subjects can be aware of this very fact. And yet, the pains still play a powerful and immutable role in motivating and guiding behavior. Unlike a visual illusion, where we can bracket off the illusory input from our all-things-considered beliefs and actions, the person with chronic neuropathy is unable to reconcile the opposing forces acting upon her. At once she feels the immense pain, while also knowing that it is not a meaningful signal, not serving any larger purpose beyond the pain itself. This disconnect, we believe, manifests itself in the torrent of harmful psychological and physiological reactions that are known to affect chronic pain sufferers, including depression, immune deficits, and other disorders.

And this is to say that, at least for creatures in whom suffering gives rise to the disconnect we’ve identified, first order suffering can produce yet further, second order suffering. Moreover, and like the first-order suffering, this second-order suffering is also unlikely to be assuaged by more, or more conclusive, rational argument: it’s hard to see how supplementing rationality in these ways could dampen our reaction to first-order suffering, given that first-order suffering is not under rational control. It is this set of well-known secondary reactions that

---

8 Pain from neuropathy is by no means unique in this respect. Many pains are caused by conditions that have little immediate health impact other than the pain.
9 Note that this has a similar structure to compulsions. A subject suffering from severe compulsion might recognize the meaninglessness of the compulsive act, and its all-things-considered emptiness as a reason, and nonetheless feel compelled to carry out the act. That compulsion exhibits this structure is, we suggest, a reason for thinking of compulsion, despite the typical absence of concomitant pain, as a species of human suffering.
10 This consideration invites worry about a looming regress: if reflection on the resistance of nth-order suffering to rationality causes n + 1th order suffering, then there will be a regress if reflection reveals the latter to be itself resistant to rationality. Luckily (from the point of view of avoiding regress), there are limits to ordinary reflection: it would seem
constitute (one variant of) the novel form of suffering we believe arises from our unique human perspective.

One place where there is considerable evidence of such secondary effects is in cases of chronic pain. Despite being often defined in temporal terms (e.g., as pains that persist for three or more months) what importantly distinguishes chronic from other pain is not simply a matter of time elapsed, but the close association between chronic pain and other negative states like depression, learned helplessness, and immune system weakness. But researchers have found that chronic pain also has important links to particular types of rational appraisal. Thus, Young Casey et al. (2008) report that depressive symptoms and negative beliefs (primarily constancy and permanence beliefs: constant negative thoughts about the pain, combined with the belief that the pain would never go away) are a more important factor in the development of chronic pain than were previous instances of trauma, duration of acute pain, intensity of suffering episodes, and baseline pain beliefs:

Baseline depressive symptoms and pain permanence beliefs were the most powerful predictors of chronic disability, uniquely accounting for nearly half of the variance predicted by the full model. Depressive symptoms and uncontrollability beliefs may lead to passive coping and avoidance, thereby exacerbating disability (75).

It would seem, then, that what best explains the difference between acute pains (which are often manageable) and chronic pains (which are often debilitating) are the negative beliefs a subject has about the pain.

What we have said above about rational disconnect is explanatorily useful in this context because it offers an account (not otherwise addressed in the scientific literature we are reviewing) of the source of such negative beliefs. Namely, if we are right that subjects undergoing such first-order suffering experience a direct, ongoing, and uncontrollable conflict between their overall rational considerations and intense signals of sensory pains, it is unsurprising that this should result in negative thoughts, especially of the constancy and permanence of the first-order states of suffering. For those with neuropathy, the pain signals are constantly intruding: what to eat, what to wear, where to sit, every decision runs through their pain. We can understand how this directly leads to negative thoughts about (at best) unusual to reflect on the relations between rationality and nth-order suffering for very many values of n. Wherever reflection ends, the present threat of regress ends with it.  

There is a lot to say here on “secondary affect,” and about whether animals for instance can suffer from depression and the like (the evidence here is spotty to say the least). See Price and Harkins (1992).  

A good (if now dated) discussion of these cases can be found in Hardcastle (1999). In addition, the more recent biopsychosocial conception of pain takes seriously the pervasive influence of social and psychological factors on pain experience and management (for a good introduction to this perspective, see the essays in Hadjistavropoulos and Craig (2004).
the nature and status of the pains. It is even worse in those cases where the pain is known to be giving a false signal, since it is not signally any immediate danger. If there is no damage or potential damage being signaled, then there is nothing to heal or correct in order to make it go away (the fact that central neuralgias are also immune to most forms of pain treatment also cannot help).

When suffering and other reasons come into conflict, the states of suffering are sticky: recalcitrant to our broader reasons and concerns, constantly intruding on our thoughts, impeding our actions, and functioning as motivating reasons despite our awareness that they are not. Sometimes — rarely — we are able to suppress the pains, and inhibit them for short durations, by convincing ourselves that the pains are actually good for us, or at least serve some positive purpose. But this requires extreme effort, and we are rarely up to the task, especially when the pains are severe or prolonged. This disconnect, we propose, leads to increased traumatic stress reactions, prevalence of negative beliefs, and feelings of helplessness and depression.

We thus suggest that there is a very compelling reason for caring deeply about the identified disconnect between sensory pains and our broader rational capacities: such a disconnect is very likely a central contributing factor in novel and debilitating forms of suffering beyond those found in the initial pain state. As we'll also see, we believe this disconnect also helps illuminate the structure and nature of our overall rational system. Indeed, we believe it marks an important area of potential discontinuity between us and other less/more sophisticated organisms.

3 Alternative diagnoses

So far we have argued that there is a peculiar disconnect between suffering and rationality in our psychologies (§1), and that this can have serious harmful consequences (§2). We want to argue that this disconnect shows us something unique about suffering. Namely, we believe it shows that suffering does not fit comfortably into our broader rational capacities in the way that other reasons do, and this stems from something special about the nature of suffering. Before we come to this, however, we want to rule out several alternative putative explanations for the disconnect — viz., explanations that attempt to reduce what is going on in cases like Vaccination to some more common, and well-investigated phenomenon. As we see it, there are four principal explanatory alternatives that need to be set aside: explanations in terms of akrasia, non-intentional action, further deliberation, and failures of deliberative decisiveness. (Obviously this list is not exhaustive, but we do think these are the most important alternatives.)
3.1 Akrasia

The problem of akrasia has a long history in philosophy, and has been treated by most authors as a problem about the possibility of weakness of will\textsuperscript{13}.

Here is a typical example.

**Akrasia:** Suppose Corine judges that it would be best for her to eat a healthier diet, and she thereafter avoids excess sugars and fats. She does this because she has reasoned that this is, all things considered, what she has most reason to do. It’s not difficult to imagine that she firmly sets her will to the task, and throws out all of her sweets, orders healthy options when dining out, etc. It’s also easy to imagine that her resolve fails when faced with temptation. On an outing, she cannot resist the temptation to join when her friends all order chocolate sundaes. Even though she knows it’s not the right thing for her to do, and still, it seems, believes this, she breaks down and eats the sundae anyway.

These cases are common, and have raised philosophical worries since antiquity (hence their name). How is this case any different from the disconnect with rationality manifested in the suffering-involving case of Vaccination we considered earlier? Can we explain our inability to sit still when faced with actual (or impending) pain as nothing other than a certain kind of weakness of will\textsuperscript{14}.

We cannot. There are several respects in which the disconnect present in Vaccination comes apart from that present in Akrasia\textsuperscript{15}.

First, canonical descriptions of weakness of will involve a subject choosing freely to do B even though she judges A to be the better option. Thus, the disconnect between deliberation and action manifested in Akrasia is a single-stage problem resulting from a gap between the beliefs/judgments and the desires involved. In contrast, in Vaccination and similar suffering-involving cases, there is a further disconnect, over and above the disconnect between the first order gap between deliberation and action. After all, when the needle hits the arm, it is not as though the subject is giving in to her other desire, and thereby acting against her better judgment. Indeed, typically she is actively fighting against the alternative movement, trying tenaciously to avoid flinching. The disconnect we’re interested

\textsuperscript{13}An exception to this tradition is Holton (1999), who argues that akrasia is not a weakness of will, but an over-readiness to reconsider one’s intentions. On Holton’s view, cases of akrasia are in fact instances of the failure of deliberative decisiveness we consider (and which, we argue, are also distinct from our cases of disconnect) in §3.4. Since we’ll be considering that further alternative below, and so will eventually provide reasons against both as reductive explanations of our disconnect, our treatment of akrasia as a weakness of will is merely an expository convenience, and should not be read as taking an official stand between Holton’s and the majoritarian story qua conception of akrasia.

\textsuperscript{14}As one might expect, Beach can easily be recast as such an example.

\textsuperscript{15}In making and elaborating this claim in what follows, we do not mean to commit ourselves to any particular philosophical or psychological account of akrasia.
in occurs even when the agent does not give in, and does not even have a desire to do so.

A second difference between Vaccination and Akrasia is that the gap involved in the former but not the latter crucially implicates a reason (a mental state or desire) that is not itself reasons-responsive. In the cases of Beach and Akrasia, the alternative reasons between which we deliberate (to have a good time/to get necessary work done; to eat healthily/to enjoy a chocolate sundae with friends) are reason-responsive, and so fit more easily into the subject’s overall psychology. What makes the suffering-involving disconnect in Vaccination is so trenchant is just that at least one of our reasons — viz., that involving suffering — is to a significant extent not responsive to the rational influence of our other reasons.

A third difference between Vaccination and Akrasia builds on the second. It is plausible that the subject in Vaccination who gives in to her suffering reason against what she takes to be her all things considered interest is less blameworthy than the subject in Akrasia who gives in to her non-suffering reason against what she takes to be her all things considered interest. We typically think of subjects as responsible or blameworthy for their weakness of will, and feel guilt or shame when we ourselves exhibit this phenomenon. But we have these reactions to a significantly lesser extent in cases involving suffering. When we give in to our suffering, it is generally thought that the suffering mitigates our responsibility. Hence, action is not as freely chosen in such cases — presumably just because, as noted, suffering reasons are not themselves reasons-responsive, so significantly less under rational control.

Beyond these differences between the sort of disconnect manifest in Vaccination and that manifest in Akrasia, there is a further reason for not wanting to explain our disconnect in akratic terms. Recall that we introduced our disconnect, above, by contrasting a case where it does arise (Vaccination) from a case where it does not (Beach). Given that there is this contrast, it is reasonable to demand of a proposed reductive explanation of our disconnect that it explain why the disconnect arises in the one sort of case but not the other. But on the (plausible) assumption that non-suffering reasons like those figuring in Beach are no less susceptible to akrasia than suffering reasons such as those figuring in Vaccination, it’s doubtful that an akratic explanation of the disconnect could meet this demand: on its face, it would seem that an akratic account of the disconnect would be no less applicable to cases involving non-suffering reasons (e.g., Beach) than to cases involving suffering reasons (e.g., Vaccination).\footnote{Might the proponent of the akratic explanation answer our demand by supplementing her account with some more specific further reason why suffering reasons are especially apt targets for akrasia? Perhaps she could; but if so, then we suggest that it is that specific further reason, whatever it is, and not akrasia by itself, that explains the disconnect.}
We take these considerations to limit considerably the attractions of understanding our disconnect in terms of akrasia.

### 3.2 Non-intentional action

In arguing for a distinction between the suffering-involving disconnect in Vaccination from the disconnect in Akrasia, we contended that the former (unlike the latter) should not be seen as a rational interplay between mutually influencing reasons, in so far as suffering reasons are not reasons-responsive. But this contention invites the thought that perhaps the suffering-guided actions under consideration in suffering cases are, because not subject to reason, not intentional at all. Thus, in the case of Vaccination, for example, perhaps the flinch or movement performed when the needle hits the arm is a kind of reflex that occurs without our sanction.

It is extremely plausible that there are pain-induced reflexes that shouldn’t count as intentional actions. When we touch a hot pan, we might jerk our hand back automatically without (in some sense) deciding to perform that action. However, we don’t think the action in Vaccination and similar cases should be understood in this way. For the actions undertaken in our sorts of cases are flexible, actively performed actions that we recognize and treat as our own. They can be, and often are, controlled and managed, though sometimes imperfectly (cf. Schroeder [2005] on Tourette’s).

In fairness, there is a sense in which suffering (e.g., pain) overwhelms us in the cases we have in mind. But this point requires care. Suffering does not literally force us to move in particular ways in such cases. Rather, it overwhelms us in making us decide to act. What is forced isn’t the physical movement or action (as in a reflex), but our executive control concerning the decision to act. Despite our best efforts, we simply cannot help but to act in the ways that we do.

One way to see this point is to note that, in many cases of suffering disconnect, the actions we undertake are highly complex and deliberate. For instance, one of the authors of this paper sometimes suffers from extreme itchiness on the hand from an old chemical burn. Despite knowing that scratching and cold water in the long term make the itching worse (by irritating the skin and drying it out), this author finds the immediate relief provided by scratching and cold water

---

17To be clear, our claim is certainly not that the suffering-involving disconnect we’ve highlighted can’t be thought of as a form of akrasia in any sense at all. But we hope it is clear that even if suffering does amount to a kind akrasia, it is a special form of the phenomena, with a range of distinctive features, that therefore deserves its own explanation.

18We take no official stand here on the complicated and controversial matter of how to distinguish intentional from unintentional action. For some influential views on this matter, see [Anscombe (1957); Davidson (1963); Goldman (1970); Searle (1983); Dretske (1988); Velleman (1992); Setiya (2007)].
irresistible. The scratching actions and the walk over to the faucet to turn on the water, etc., are extended in time, controlled, and seemingly voluntary, although the executive decision to act in these ways — though they are known to be all things considered suboptimal — is not. These sorts of actions involve a kind of disconnect, but not, we claim, as a result of the actions failing to be intentional. Something similar is going on in Vaccine: if the subject decides to move because of the pain, it need not be a reflexive jerk, but could easily be a complex action involving bodily movements and verbal requests.

There’s a further reason for us not to accept the present suggestion that the suffering-guided actions under consideration in suffering cases are non-intentional. Treating the actions resulting from suffering as non-intentional because not subject to reason is at odds with the supposition that suffering is an element of the space of reasons — that it provides reasons for the actions in which it eventuates — with which we began. Giving up that supposition is not an alternative solution to our puzzle, but an abandonment of it, and would seem to fly in the face of the robust intuitive and theoretical support that, as we observed in the opening of the paper, the supposition appears to enjoy. We are, therefore, inclined to seek other avenues of response.

3.3 Spike in desire, new deliberation

Perhaps what happens in the cases of apparent suffering disconnect is that we have simply misjudged matters in our initial deliberation and come to a new and different practical conclusion that incorporates new information we hadn’t initially realized we needed.

Thus, in the case of Vaccination, we decided at time $t_1$ to stay still on the basis of the best available evidence (it was, at $t_1$, what we had most reason to do). But by $t_2$, arguably the epistemic situation had changed — we had come to know just how painful that needle was really going to be (or how much it was going to bother us, etc.), and staying still was no longer what we had most reason to do.

The idea then is that once we get the new information, once we are exposed to the feeling, we quickly (maybe very quickly) engage in a new round of practical deliberation and come to a new conclusion: we should move away from that needle, and fast (or, maybe we’re not that worried about our overall health after all, etc.). One way this could happen is if the pain state itself spiked our desire against it in a way that pushed that reason ahead of our desire for longterm health.

If this is the right way to think about this case, then again there is no real disconnect.\footnote{The case is, also, importantly different from cases of OCD or addiction, though in these cases, too, we might also find the very form of disconnect we’re interested in; see note 9.}

\footnote{In conversation, Rashida Ahmad suggested that perhaps there is only a difference here between occurrent versus predictive acts of deliberation. In the pain case, perhaps the}
But there are reasons for believing that the suffering disconnect is not best understood in terms of such a spike in desire and new round of deliberation. Crucially, if there really were new deliberation, then ceteris paribus the subject should, after the flinch, stick with and accept the conclusion of that later round of deliberation — which means sticking with the rejection of the conclusion of the initial round of deliberation. But, though there are cases that take that form, that doesn’t seem to be the general description of our sorts of suffering cases. In many cases, including a typical instance of Vaccination, the subject does not disavow the original practical conclusion. Rather, she persists in believing that the choice she had initially made remains all things considered best. She does not prefer the deviating action. Indeed, a common occurrence is for someone to flinch once or twice, but to keep trying again and again to fulfill the original settled decision. This does not seem consistent with the pattern of reassessment and new deliberation.

3.4 Failure of deliberative decisiveness

A yet further possible redescription construes the alleged disconnect in suffering cases as a failure of deliberative decisiveness. Here the story would be that, contrary to our initial description of the cases of Beach and Vaccination as alike from the point of view of rationality, there is a characteristic failure of rational decisiveness in Vaccination and other cases involving suffering (and not elsewhere) which has the result that suffering-involving lesser reasons (such as $R^*_{\neg}$ in Vaccination, but unlike $R^-$ in Beach) don’t cease to be action-guiding. This failure comes about because practical deliberation becomes Hamlet-like, unable to come to the point of a decisive verdict about how to act, and instead continuing to oscillate between reasons when it is applied to suffering-involving reasons.

Unfortunately, there are two serious defects with the Hamlet proposal, when construed as a diagnosis of the asymmetry.

The first is that it only pushes back, but does not resolve, the explanatory question posed by the disconnect we have observed in suffering-involving reasons. Whereas we began by asking what it is about the suffering-involving reason $R^*$ that (unlike the non-suffering-involving reason $R^*$) makes it retain its action-guiding force despite the exercise of rationality, adopting the Hamlet diagnosis trades that initial question for the new question of what it is about the suffering-motivating reason is only present when the pain is occurrent. When we are not in pain, and yet deliberating about what to do in some future state involving pain, we are comparing apples with oranges.

We do not think this is plausible. First, our disconnect happens even when the pain and the alternative are both occurrent (as in Vaccine). Second, if this were the right diagnosis for the disconnect, we should expect non-suffering versions of the disconnect trading only on the occurrent-non-occurrent distinction. We cannot think of any.

21 Thanks to Fiona Macpherson for pressing us to consider this possibility.
involving reason \( R^{\ast\ast} \) that (unlike the non-suffering-involving reason \( R^{\ast} \)) permits practical rationality to reach a decisive terminus when applied to it. Since we don’t see the second question as any easier than the first, we don’t see that much has been gained in passing from the one to the other.

The second defect of the Hamlet proposal is that it appears to have dubious entailments. Specifically, because it locates the asymmetry in a difference of the efficacy of practical reason, the Hamlet proposal would seem to predict that the asymmetry would be lessened or eliminated in subjects suffering from deficits of practical reasoning (e.g., juveniles, psychopaths) who (because of their deficits) are less successful at engaging their practical reason even in cases where their reasons are not suffering-involving. Because such subjects are less likely to engage the processes that, on the proposed diagnosis, risk turning asymmetrically Hamletian just when suffering is involved, we should expect not to see an asymmetry between their responses to suffering-involving and non-suffering-involving reasons. Alas, there is no reason to believe that this prediction is correct.  

4 Ambivalence though partial integration

The shortcomings of the options considered in §3 suggest strongly that the disconnect we’ve highlighted between states of suffering and rationality cannot be adequately accounted for by reducing it to other well-known forms of rational disconnect. If so, the disconnect we have identified is a distinct syndrome with its own psychological profile. In this section we attempt to characterize that profile, and then show why it reveals something peculiar about our kinds of minds.

---

22In fairness, there is some evidence that suffering may be decreased in certain populations — e.g., subjects with attention deficit hyperactivity disorder, or patients with frontal lobe damage (including damage resulting from lobotomies) of the kind described by [Damasio (1994)] — that are canonically described as involving deficits in practical reasoning. But we take this point not to undercut what we say in the main text.

First, the deficits in these subjects appear to result from excessively Hamletian practical reasoning: these are not cases where subjects fail to engage a system that risks turning Hamletian, but cases where subjects do engage a system that does turn Hamletian (unusually often). As such, they are not the kind of subjects who fall under the prediction of the Hamlet diagnosis to which we are objecting. And, second, even an observed symmetry in these subjects’ responses might not reflect a genuine symmetry in their reactions to suffering-involving and non-suffering-involving reasons: for there is reason to believe that such subjects experience less suffering than others [Damasio (1994), 60], hence have fewer or less powerful suffering-involving reasons.
4.1 Partial integration

Reflection on what is shared and not shared between cases like Beach and cases like Vaccination suggests that the disconnect concerning suffering is plausibly not understood as a difference in the operation of rationality.

On the one hand, as we noted in characterizing the cases above, the operation of rationality in both cases appears to be relevantly similar: in both cases, rationality succeeds in converting negative reasons ($R^−$ in Beach, $R^∗−$ in Vaccination) from being initially potentially action-guiding to being no longer action-guiding. On the other hand, as we also noted earlier, negative reasons in suffering cases appear to withstand the operation of rational deliberation, and to persist in exerting control over our actions despite the operation of rationality, in a way that negative reasons in non-suffering cases do not. To say this is to note that, while the psychological role of non-suffering negative reasons is (to a significant extent) under rational control, there is an element of suffering negative reasons that falls outside rational control. This is why, once the initial psychological force of $R^−$ is mitigated by the exercise of rationality, we no longer have a negatively valenced reaction against $R^−$. Whereas, in contrast, even after the initial psychological force of $R^∗−$ is mitigated by the exercise of rationality, we nonetheless continue to have a negatively valenced reaction against $R^∗−$.

This pattern of facts suggests that, while negative reasons of both suffering and non-suffering kinds are to a considerable extent integrated with and responsive to rationality, suffering negative reasons are only partially integrated with rationality. They are not wholly arational, or outside the space of reasons altogether: they are, after all, reasons. But, unlike non-suffering negative reasons, they appear to involve two factors — one responsive to rational deliberation and the other not. In this sense, what the disconnect shows is that suffering negative reasons are partly rational and partly arational, and in this way only partially integrated with rationality.\footnote{Dialectical note: Our claim that there is a component in suffering reasons that resists rational control is intended as a descriptive proposal, rather than a reductive analysis. What we have said leaves open the interesting question \textit{why} suffering reasons have a residue that resists rational control. That it leaves this question unanswered would be a shortcoming if, like the suggestions canvassed in \cite{3}, our claim were offered as an attempt to explain away the disconnect reductively in terms of other phenomena.}

4.2 Suffering in minds like ours

If we are right that the rational disconnect observed in suffering reasons is a result of the dual, partly rational and partly arational, character of those reasons, then this suggests that the disconnect should only arise in creatures like us, whose psychologies exert both rational and arational pulls on states of suffering.
Of course, suffering can play a powerful psychological (and, presumably, adaptively advantageous) role in simple creatures just as it can for us. But the present thought is that the powerful psychological role suffering plays in simple creatures need not run through rationality. Thus, though it is plausible that (e.g.,) fish/cats/mice avoid sensory pains, it is unclear that such creatures would endure sensory pains that they recognized as in the service of some overall good reason, or that they would be more upset about pains that they deemed not in the service of some overall good reason. If this is right, it is tempting to think of the psychological role of suffering in such creatures as running through only one of the two distinct pathways that are operative in us. But if there is only this single pathway, there is no possibility of conflict, hence no possibility of the kind of ambivalence to which we find ourselves vulnerable. Suffering will be simply, straightforwardly, and unambivalently motivational for such creatures, as it is not for us.

On the other hand, it is also possible to imagine a kind of hyper-rational creature in whom suffering-involving reasons are treated exactly like non-suffering reasons in being completely under rational control. This kind of creature would cease to have a negative reaction to suffering-involving reasons like that in Vaccination once the latter had been mitigated by the successful exercise of rationality. Plausibly, suffering would be motivational in these hyper-rational creatures in just the way (whatever that is) that non-suffering reasons can be motivational. But, again, this relegates the motivational force of suffering-involving reasons in such creatures to just one of the dual pathways available in our own psychologies, and so removes the possibility of the kind of clash to which we are susceptible. Again, suffering will be simply, straightforwardly, and unambivalently motivational for such creatures, as it is not for us.

The peculiar ambivalence toward suffering that we experience can only arise because our psychologies combine both the lower-level systems for responding to suffering present in simpler creatures and the flexible, rational system for

---

24 These points need to be treated separately from cases of classical conditioning, in which even very simple organisms can be trained to endure a small shock in order to receive a reward. Such conditioning can be explained through simple causal learning mechanisms that require no rationality.

25 That said, there could be a distinct kind of conflict between reactions to forms of first order suffering (say, between reactions to thermal pain and thirst) that could arise even in the absence of reflection and reasoning, hence in simpler creatures in whom the psychological role of suffering does not run through rationality. We take this point not as contravening our contention that suffering is unambivalently motivational (in our sense) for such creatures, but as showing that they may have multiple unambivalently motivational states that pull in different, and possibly opposed, ways. Indeed, that is plausibly true both for these simpler creatures and for more psychologically complicated creatures, such as ourselves, who are additionally susceptible to the particular kind of ambivalence about suffering that we have highlighted.
responding to suffering present in the imagined hyper-rational creatures, and because these distinct systems can come into conflict. Moreover, and significantly, when they arise, these conflicts cannot be assuaged by rational argument. You may have good reasons for believing that an episode of suffering is all things considered worth undergoing for some greater good, and even that it is not associated with any significant harm. But none of this makes the suffering go away.

That such conflicts can arise and persist in creatures like us, with highly flexible cognitive systems with capacities for introspection and self-reflection, means that the automatic functioning of our (generally low-level) suffering system is for us not just an adaptive key to survival (surely it is, as in the rest of the animal kingdom) but also, and additionally, a point of worry and social anxiety caused by its unsteady connection to reason. In this sense, the role of sensory and social suffering for us, while adaptively connected with motivation and reason, is different from — and, as we have argued in §2, carries significant further negative consequences for well-being relative to — the role it has in either simpler creatures who only respond to occurrent negative stimuli or in imagined hyper-rational creatures who can reason away their suffering.

5 Conclusion

That our reactions to suffering embody conflict, and hence ambivalence, in the way we have been trying to bring out, is, it seems to us, an interesting feature of our mental lives. Of course, the idea that there is a conflict between rational and arational elements of our psychologies is a standard trope of modernity. For this reason alone, it would not be surprising to find some sort of conflict between rationality and suffering falling out of views on which the latter fails to provide reasons. What we have been suggesting, however, is that psychological conflict arises as well on the opposite supposition that suffering does provide reasons. Indeed, the conflict is in a way more pressing on the latter view. For a conflict between reasons and non-reasons might seem to be a clash between elements too disparate in kind to be mutually constraining — perhaps in that case the best response would be to resort to a coin-flip. In contrast, if, as we are suggesting, the conflict arises within the space of reasons, it is to that extent easier to see the elements as commensurable and mutually constraining. Perhaps this explains why the conflict we experience in such cases is such a powerful source of anxiety.\footnote{We are grateful to Marilyn McCord Adams, David Bain, Brock Bastian, Michael Brady, Jennifer Corns, Ross Hetherington, Tom Johnstone, Fiona Macpherson, Alisa Madrigal, Manolo Martinez, and Abraham Sapién Córdoba for discussion and comments on earlier versions of this material.}
References


20


