

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 us. It is overwhelmingly plausible from the point of view of both naïve reflection and broadly empirical investigation that sensory/psychological 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 hard with a

*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 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.)

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 motivating *reasons* to act in certain ways. (For Lucy perhaps pain gives a reason to tend to her injured thumb; for Luke perhaps the suffering caused by 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 durable ways, even in cases where you might have expected the two to come apart. Thus, for example, it is interesting that those incapable of physical suffering (e.g., 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 genuine 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).² 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).³

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.⁴ However, accepting this picture invites a concern that, to our knowledge, hasn't been discussed elsewhere, over the nature

²Interestingly, 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. This reflects a crucial ambiguity in the motivating role of suffering itself to which we will return in §4.

³Similar effects have been found in humans; see e.g., the discussion of the human pain inhibitory system in Hardcastle (1999, 132–134).

⁴In 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, 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 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).) One might also worry about whether the category of suffering is sufficiently unified to support any interesting and generalizations at all. Indeed, some have argued that the heterogeneity of sensory pain threatens the prospects for systematic theorizing about pain (Heathwood 2006; Smuts 2010); and those worries are

of and limits to the interaction between suffering and rationality. This concern will be the focus of this paper.

One way to bring out the disconnect 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*. Or at least, if they are reason-responsive, then they are reason-responsive to a more limited extent than other of our reason-giving states. We seem rationally motivated by our states of suffering, but our states of suffering do not seem sensitive to our broader rational circumstances in the same way and to the same extent.

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 (again, depending somewhat on one's account of reasons). 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.

If pointing to this were all we had in mind, the observation would be of limited interest. But the case of suffering is plausibly different from and more interesting than the case of the water's being in the fridge (etc.) in at least two respects. First, the rational role of suffering for subjects like Lucy and Luke is significantly more extensive than that of water's being in the fridge for Theo. Second, there is a sense in which states of pain and suffering are both internal to and part of a subject's own experience. In cases where Theo is unaware that there is water in the fridge, there's a good sense in which that fact, though a reason for him, is no part of his psychological economy. In contrast, there is no question that felt pains and other states of suffering are psychologically available to the subjects in whom they occur. They are, like subjects' beliefs and desires, a central part of their current psychological profiles. And yet, like things typically outside of our psychologies, states of suffering seem somehow resistant to our broader rational concerns. While we obviously don't have complete control over all of our mental states and experiences, suffering is among the most salient and intense of those states, and plays a more significant motivating role than other typically uncontrolled mental states. It hardly needs to be said that states of suffering play a role far more central than other typical external entities (which, after all, on most views get their motivating force only after uptake into a subject's psychological economy).⁵

presumably more pressing as applied to the yet (far) more inclusive category of suffering. (We return to this worry in §4.3.)

⁵In fact, states of suffering may not be unique in this respect even among states that are internal to subjects' psychologies: arguably some perceptual states and non-instrumental

The claim with which we began, 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 very fact that makes 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 in critical ways with practical rationality in creatures like us, there seem nonetheless to be significant limits to that integration. In particular, when suffering becomes (practical) reason-conferring for a subject (when it enters into what Sellars (1956) famously calls the “space of reasons”), it nonetheless continues 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, outside influence. In this sense, suffering seems to play a dual (partly rational, partly not rational) role in the mental lives of creatures like us. Interestingly, and as we shall argue below, this duality seems peculiar to our sorts of mental lives: it likely does not occur in the mental lives of psychologically less or more rationally sophisticated creatures.

Before we move on, it will be useful to limit the scope of discussion by mentioning (so as to set aside) a couple of aims that we do not have in the present paper. First, while we take the disconnect between reason and suffering that we are highlighting to be interesting and worthy of discussion (partly because of the surprising predictions and explanations that, we argue below, it makes available), we do not offer it as a necessary or sufficient condition for the occurrence of suffering. Indeed, our position is that the disconnect is itself a *further* consequence — and one that can have serious negative effects on the creatures in whom it arises — of an antecedent state of primary suffering for which we have no analysis to offer in the present paper. Second, nothing we say here is an attempt to explain the inherent *badness* of such suffering, nor to say anything directly about how states of suffering motivate our behavior. We set aside for now the issue of what makes suffering bad (but see Aydede and Fulkerson 2015), and assume that suffering has an antecedent negative valence. Our concern is to explore the nature and consequences of the disconnect that, we claim, arises for states that have this negative valence.

Having so qualified our aims, here is our plan for what follows. We’ll begin our discussion by attempting to more precisely characterize the disconnect discussed above (§1), and arguing that it has serious negative psychological consequences (§2). Next we’ll ask how best to understand this 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).

desires are reason-conferring but not reason-responsive. (Thanks to David Bain for this observation.)

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, and you don't get an extension on your submission), the rational thing for you to do would be to satisfy that stronger 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. The pull that it exerts is sensitive to your other concerns and decisions, such that your decision to act on the basis of one consideration seemingly weakens the pull of your reasons. 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 (or if it is, that activity is influenced by and sensitive to the reasons working in the opposite direction). 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 the patient remain completely still for two minutes while the drug is administered via an intravenous needle. If the patient moves 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^{*-} 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.

That analogy notwithstanding, it seems that there is an important difference between the cases. Namely, in the case of Vaccination, unlike the case of Beach, where the reasons did not involve suffering, there is a way in which R^{*-} persists, or resists being overwhelmed by rational control. In this way, R^{*-} continues to exert control over our actions even after rational deliberation has done its work.⁶ 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.⁷

⁶Michael 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.

⁷A 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).

It would appear that the action-guiding force of the suffering-involving reason R^{*-} in Vaccination survives the exercise of rationality — and in this sense is interestingly disconnected from many of our broader reasons — 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 endured pains and suffering that continued to impact our actions despite strong rational considerations acting in the opposite direction. An ordinary example of this kind involves spicy food. Capsaicin works by binding at relatively low temperatures to TRPV1 receptors in the nose, mouth, and skin that are ordinarily activated only by high temperatures; it thereby tricks 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, or, more importantly, to reduce its motivational force.⁸ In what follows, we want to argue that this disconnect holds significant lessons about the nature and role of suffering in the mental lives of creatures like ourselves.⁹

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 limited, effortful, and temporary, we are inclined to treat them as exceptions that prove our more general rule.)

⁸A 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).)

⁹Objection (pressed on us by Dana Nelkin): Our presentation of the disconnect, which rests only on an intuitive understanding of pain and a few examples (as opposed to any systematic account of pain) may crucially conflate features of pain with those of numerically distinct states or processes. If, for example, it is not pain but some concomitant state that is reason-conferring, then there is no puzzling disconnect, for there is no one entity that is both reason-conferring and reason-resistant.

Response: First, just to be clear, the puzzle arises for states of suffering generally, rather than exclusively for pains, and this point already makes it somewhat unlikely that the puzzle depends essentially on differences in the theoretical understanding of the subclass of pains. Moreover, as we have argued elsewhere (Cohen and Fulkerson 2013), proponents of different accounts of pain (say, attitudinal theories (Feldman 2002; Heathwood 2006), evaluationist theories (Helm 2002; O’Sullivan and Schroer 2012; Bain 2013), functionalist theories (Aydede 2014; Clark 2005; Fulkerson 2014), and imperativist theories (Klein 2007; Hall 2008; Martínez 2011)) can and should accept the claim that pains confer reasons as a *prima facie* datum which it is incumbent on their theories to explain. If so, any adequate account will have to cite some theoretical element (whether numerically identical to the

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 sensory and psychological suffering, 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 in §4.

Human suffering is often rooted in bodily reactions, but it is also 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 how much pain we feel; social context can influence pain judgments; and will power is inversely correlated with pain tolerance (Ossipov *et al.* 2010; Silvestrini and Rainville 2013).

One area of 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 central and peripheral neuropathy. 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.¹⁰ 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 relatively immutable role in motivating and guiding behavior.¹¹ At once she feels the immense pain, while

pains themselves or not) that makes pains reason-conferring, and with respect to which a version of our puzzling disconnect arises. We can ask: why is it that that element (whatever it is) makes the pains with which it is connected reason-conferring and extensively integrated with rationality, but does not also make those pains susceptible to rational mitigation in the way that non-suffering reasons seem to be?

¹⁰Neuropathic pain is by no means unique in this respect; many pains are caused by conditions that have little immediate health impact other than the pain.

¹¹Note that this has a structure similar to that of 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.

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, this primary suffering can produce yet further, secondary suffering. Moreover, and like the primary suffering, secondary 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 the suffering, given that primary suffering is not under rational control.¹² It is this set of well-known secondary reactions that constitute (one variant of) the novel form of suffering we believe arises from our peculiarly human perspective.

One place where there is considerable evidence of just such secondary effects is in cases of chronic pain.¹³ Despite being often defined in temporal terms (e.g., as pains that persist for three or more months) what critically 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 significant 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).

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. (Cf. McKay and Dennett (2009, 499-500).)

¹²This consideration invites worry about a looming regress: if reflection on the resistance of n -ary suffering to rationality causes $n + 1$ -ary 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 (at best) unusual to reflect on the relations between rationality and n -ary suffering for very many values of n . Wherever reflection ends, the present threat of regress ends with it.

¹³A 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).

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 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 primary suffering experience a direct, ongoing, and uncontrollable conflict between their overall rational considerations and the practical influence of the intense signals of sensory pains, it is unsurprising that this should result in negative thoughts, especially of the constancy and permanence of the primary 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. This is because they have painful experiences, and the strong motivation to remove these experiences is part of what it is to have these experiences at all. No outside reason will mitigate this form of motivation unless it also removes or lessens the experiential qualities associated with the pains (which is why, for instance, being distracted or attending to other stimuli can indeed have a positive impact on suffering episodes). We can understand how this directly leads to negative thoughts about 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 signaling 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). We are instead left only with the self-motivating experience.

When suffering and our 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 ought 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. This often weakens or pushes to the side the very experiential qualities motivating us in the first place (as above). 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 compelling reason for caring deeply about the identified disconnect between sensory pains and our broader rational capacities: such a disconnect is likely a central contributing factor in novel and debilitating forms of suffering beyond those found in the initial pain state.

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). Before we go on to offer our own characterization of the disconnect in terms of the way that suffering fails to fit comfortably into our broader rational capacities as other reasons do (§4), we want to rule out several 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 five principal explanatory alternatives that need to be set aside: explanations in terms of akrasia, non-intentional action, alief, further deliberation, and failures of deliberative decisiveness. Obviously this list is not exhaustive, but we 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.¹⁴

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

¹⁴An 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 §3.5. 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 can be treated as merely an expository convenience, rather than as our taking any official stand between Holton's and the majoritarian story qua conception of akrasia (or, for that matter, as committing to any particular account of akrasia).

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?¹⁵

We cannot. There are several respects in which the disconnect present in Vaccination comes apart from that present in Akrasia, at least as the latter is usually understood.

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 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 that is not itself responsive to our broader reasons. 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 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 reason-responsive, so significantly less under rational control.¹⁶

¹⁵As one might expect, Beach can easily be recast as such an example.

¹⁶Having distinguished in these ways between the kind of disconnect present in Vaccination and the kind present in Akrasia, we note that it may, nonetheless, be possible to describe particular cases — perhaps even the sundae-eating case we labeled 'Akrasia' — on the latter model. On this description, overeating is more like acting under compulsion than

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 akasia 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).¹⁷

We take these considerations to limit considerably the attractions of understanding our disconnect in terms of akasia.¹⁸

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 typically responsive to our broader reasons. But this contention invites the thought that perhaps the suffering-guided actions under consideration in suffering cases are, because only to a limited extent subject to reason, not intentional at all. Thus, in the case of Vaccination, for example, perhaps

a weak-willed choice, and patterns with Vaccination rather than Akasia on the dimensions just considered: it involves fighting against rather than giving in to an opposed reason; the triumphant reason is not reason-responsive; and because it is not fully under rational control, our succumbing to it seems less apt for blame and responsibility.

But the availability of a non-akratic/compulsion understanding of this and other cases traditionally labeled as akratic does not undermine our distinction between the disconnect present in Vaccination and the disconnect present in Akasia. Rather, it reinforces our point that there are two quite different models here — and reveals that there is controversy about which of them is the best description of individual cases.

¹⁷Might 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 akasia? Perhaps she could; but if so, then we suggest that it is that specific further reason, whatever it is, and not akasia itself, that explains the disconnect.

¹⁸To be clear, our claim is certainly not that the suffering-involving disconnect we've highlighted can't be thought of as a form of akasia in any sense at all. But we hope it is clear that even if suffering does amount to a kind akasia, it is a special form of the phenomena, with a range of distinctive features, that therefore deserves its own explanation.

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 these non-reflexive cases. 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. This is true even of slight pains, which, like more severe pains, do not force us to move our bodies, but are still a form of suffering and still have motivating force. They have this force, as we've said, because of the way they (and other pains) are presented to us. In severe cases, these experiential pains can become decisive. Though, in such cases, we cannot help deciding to act, the flexibility exhibited by the ensuing actions suggests that the latter are intentional in ways that reflexes are paradigmatically not.¹⁹

Another 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 sometimes finds the immediate relief provided by scratching and cold water 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.²⁰ 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 Vaccination: 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.

¹⁹We 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).

²⁰The 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 11.

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 at the outset, the supposition appears to enjoy. We are, therefore, inclined to seek other avenues of response.

3.3 Alief

In a pair of influential papers, Gendler (2008a,b) proposes the existence of a cognitive state of “aliefs” — roughly, tacit and automatic belief-like states that guide behavior despite being at odds with both rational norms and the explicit beliefs of the subject in whom they occur. Some simple cases from Gendler's many putative examples of aliefs are: (i) the state of the man suspended safely in an iron cage above a cliff that causes him to tremble, (ii) the state of the person who has knowingly and deliberately set her watch five minutes fast that causes her to rush, even though she knows that her watch is inaccurate, and (iii) the state that causes a hungry person to shrink from eating fudge shaped like dog feces (Gendler 2008b). While there is much to say about both the proposed examples and the proposed category of alief, the notion is important in the present context because aliefs appear to share the crucial features of states of suffering we are pointing to: they are motivating and guide action, yet seem relatively immune to our broader rational concerns. Rational argument (whether supplied by the subject or another) is singularly ineffective against the trembling man's trembling, the watch advancer's rushing, and the fudge eater's reluctance: the relevant aliefs persist despite countervailing rational considerations, and continue to guide action in directions subjects know goes against their best interests. In short, then, it appears that aliefs display something closely akin to the rational disconnect we have claimed attends states of suffering. It may be tempting, therefore, to suggest that our suffering disconnect can be reductively explained in terms of alief.

Unfortunately, however, we don't see how such a reductive explanation could go. For, on the one hand, states of suffering can't be *identical with* aliefs: among other differences, states of suffering are (especially) consciously accessible and negatively valenced, while, ordinarily, aliefs are neither. And on the other hand, it won't help to hold that states of suffering are merely *associated with* (/give rise to) numerically distinct aliefs, such that it is not the state of suffering, but the concomitant alief, that is the true locus of rational disconnect. For, even assuming there were such an association between suffering and alief, it remains true that states of suffering themselves (whether or not this is also true of any concomitant states) *do* elude

rational control. As such, we cannot envisage any promising way of explaining away our disconnect in terms of alief.

3.4 Spike in desire, new deliberation

Perhaps what happens in the cases of apparent suffering disconnect is that there are really two separate steps of deliberation, and each is internally consistent. The idea is that we may have 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, perhaps the best description of the case of Vaccination takes this form. Perhaps 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 , perhaps 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. So we engaged in a second, distinct act of deliberation, and, based on the new evidence, simply changed our mind in a rational way.

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.). If so, then there isn't anything like partial integration or a distinct kind of reason, and there is no real disconnect to be accounted for; instead there are just two distinct evidential circumstances and an ordinary, rational change in our decisions based on the new evidence.²¹

There are reasons for believing that the suffering disconnect we've pointed to is not best understood as involving two distinct acts of deliberation with different evidential circumstances. Crucially, if there really were new, separate deliberation, then *ceteris paribus* the subject should, after feeling the pain, 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 cases like Vaccination, the subject does

²¹In conversation, Rashida Ahmad has suggested that perhaps there is only a difference here between occurrent versus predictive acts of deliberation. This is partially right, in the pain case, as we've said: perhaps part of the motivating reason is only present when the pain is occurrent. But our puzzle can't be fully accounted for in terms of an occurrent/predictive difference. First, our disconnect happens even when the suffering and the alternative are both occurrent (as in Vaccination). 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.

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 or change her mind or succeed in undermining her other rational concerns. 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 “fighting through” is an important marker of the kind of disconnect we’ve been highlighting. None of these features are consistent with the pattern of reassessment and new deliberation suggested by this alternative.

3.5 Failure of deliberative decisiveness

A yet further possible redescription construes the alleged disconnect in suffering cases as a complete failure of deliberative decisiveness.²² 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^{*-} 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 continues to oscillate between reasons when it is applied to suffering-involving reasons.²³

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-involving reason R^{*-} that (unlike the non-suffering-involving reason R^*) permits practical rationality to reach a decisive terminus when applied to it. Since we don’t

²²Thanks to Fiona Macpherson for pressing us to consider this possibility.

²³Another possibility might fall between the alternatives considered in §3.4 and §3.5 (many thanks to Dana Nelkin for suggesting this possibility). On this alternative, a subject might have multiple, alternating spikes in deliberation, leading to something that looks like a failure of deliberation, even though at any one time the subject is completely committed to one alternative or the other.

We do not think this proposal is a live option. For one thing, it cannot be applied to cases of disconnect that involve a sustained, occurrent state of suffering. When we exert mental effort to overcome the pain in Vaccination, we are not changing back and forth between the various options. Second, it is difficult to make rational sense of a psychology that endogenously switches between alternative options in this way: it would be very unstable, at the very least.

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.²⁴

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, better 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, argue that it reveals something peculiar about our kinds of minds, and address remaining taxonomic worries.

4.1 Partial integration and experiential motivation

We have claimed that states of suffering are, while sources of rational motivation, only partially connected with our rationality. But we now want to refine this claim.

²⁴In 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.

As we see it, states of suffering are rationally motivating in two distinct ways — and only one of these forms of rational motivation gives rise to our disconnect.

What are these two forms of motivation that attach to states of suffering? First, when we undergo pains or other states of suffering, we are rationally motivated by these states to reduce, overcome, or eliminate the underlying distal cause of our suffering. Thus, for example, when you undergo a pain caused by touching something hot, you are rationally motivated to stop touching whatever it was that caused the pain. But, additionally, when we undergo pains or other states of suffering, we are rationally motivated by these states to reduce or eliminate the suffering experience itself — even if we know that doing so does nothing at all to correct or remove the underlying causes of the pain or other suffering. This is why, for example, we often select, use, and regard as effective medical treatments for pain that reduce our experiential suffering without treating or doing anything to address the underlying cause of the suffering.²⁵ Thus, states of suffering, in a way that distinguishes them from other reason-providing states, partly motivate us to act in a way directed on our very experiential states themselves. They signal reasons for us to act that are partially generated by, and directed on, their own experiential qualities.²⁶ In so doing, such experiential states bring along their own reasons, simply by their very presence. They are, as we might say, *experience-directed motivators*.²⁷

²⁵This is not always the case, of course. Many analgesics also help with the underlying conditions. Ibuprofen, for instance, reduces inflammation which is often a contributory cause of pains (especially muscle aches). Still, we take the ibuprofen primarily because it alleviates the painful experience, and only indirectly because it reduces inflammation.

²⁶Though we say that this feature distinguishes states of suffering from other reason-providing states, there is room for argument here about just how far it extends. Thus, while the feature at issue does not seem to be exemplified by ordinary doxastic states such as beliefs, it does seem sometimes seem to attach to desires. Suppose you strongly desire chocolate milk, and therefore have a pro tanto reason to acquire chocolate milk. But suppose the doctor has ordered you not to satisfy this desire; the sugar is too dangerous to your health. In this case, you now might have an additional reason directed at the desire itself. Having a strong desire that cannot be satisfied now gives you a reason to remove the desire itself. Interestingly, strong unfulfilled desires are deeply unpleasant (e.g., in the context of romantic love); and their sharing this feature might therefore provide a principled reason for counting them as states of suffering. (Reasons of space prevent us from taking up here issues about the connections between unpleasant desires and suffering; we hope to return to these matters on another occasion.)

²⁷We hasten to add that nothing about the claim that states of suffering are experience-directed motivators is intended to commit to a particular metaphysical account of pains or suffering. On the contrary, most theories of pains agree that pains are intrinsically motivating, and often give us reasons to act against their very presence or experiential qualities (see Cutter and Tye (2011) for the rare view that denies such experientially directed reasons, and recent debates about whether it is rational to “kill the messenger” (Cutter and Tye 2011; Bain 2013; O’Sullivan and Schroer 2012).

Noting that states of suffering are experience-directed motivators in this sense allows us to locate more precisely the source of the rational disconnect that arises for these states. Namely, we want to suggest that the disconnect is a feature of the experience-directed motivation, rather than the non-experience-directed motivation, provided by states of suffering.

To see this, consider poor Lucy once again, who now feels a sharp, consistent pain in her leg.

On the one hand, the non-experience-directed motivating reason the pain provides for Lucy — say, a reason to visit the doctor and undergo treatment to address the underlying cause of the pain — seems reasonably well-integrated with (i.e., not disconnected from) her rationality. This non-experience-directed motivating reason arises because Lucy's state of suffering, like any state that plays a role in rational practical reflection, provides her with actionable information about the distal world. Lucy's pain enables her to note that there is something amiss, and to undertake rationally well-advised, practical steps to address it. Significantly, and as in other cases of non-suffering practical deliberation like Beach, her state of suffering provides reasons and motivates in ways that can easily influence and be influenced by her broader, objective reasons. New information or evidence, or competing objective reasons, can influence, boost, or suppress her suffering-provided non-experience-directed motivation, just as they do in non-suffering cases like Beach. In short, the non-experience-directed motivation provided by Lucy's suffering in this case seems no less fully integrated with her broader rational life than are motivations in typical, non-suffering, cases of practical rational reflection.

On the other hand, however, there is an experience-directed motivating part of the suffering that does not integrate in this way with Lucy's broader reasons. *It doesn't matter* what Lucy's all-things-considered reasons might be, what her plans for the day might be, or what she would rather be doing. When she feels the sharp, intense pain in her leg, she is immediately motivated to alleviate *it* (i.e., the *experience*) in some way. This motivating reason, directed at the suffering itself, is not as easily and fully sensitive to Lucy's other reasons. Suppose Lucy is informed by her (trustworthy, knowledgeable, and reliable) doctor that there is nothing physically wrong with her leg, nothing amiss that needs to be addressed or treated. Even so, she still has reason to act, she is still motivated to alleviate the painful experience. So long as she undergoes it, her painful experience is likely to retain its status as something that provides her with this second type of experience-directed rational motivation.²⁸

The observations that states of suffering are motivating in both experience-directed and non-experience-directed ways, and that the experience-directed motivations they provide (as opposed to the non-experience-directed motivations

²⁸Of course, as we've noted, there can be top-down influences on many forms of suffering. But the fact that these potential influences are difficult to induce and sustain, cognitively draining, and limited in effectiveness supports our claim in the main text.

they provide) are often imperfectly integrated with broader rationality do not, by themselves, go terribly far in explaining the original disconnect between suffering and rationality. For they leave open both why states of suffering have this dual structure and why experience-directed motivators fail to integrate smoothly with our broader rational concerns.²⁹

However, our observations do allow a more specific description of the initial disconnect. Namely, they allow us to say that the rational disconnect exhibited by states of suffering traces to their associated experience-directed motivations (a kind of motivation that is connected with their felt nature), as opposed to their associated non-experience-directed motivations (a kind of motivation that is connected with their informativeness about their distal causes). Suffering without a felt component would not, it seems, be rationally disconnected in the way that it seems to be.

4.2 Suffering in minds like ours

If we are right that the rational disconnect observed in states of suffering is a result of the dual, partly rationally integrated and partly rationally non-integrated character of the motivations they provide, then this suggests that the disconnect should only arise in creatures whose psychologies exert both rationally integrated and rationally non-integrated pulls on states of suffering.

Of course, suffering plays 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. When a simpler creature feels pain it is only motivated by the experience-directed motivating aspects of the experience. As far as we know, rabbits and birds and fish, for instance, don't endure pain and suffering on the basis of broader rational reflection. Indeed, part of what we take their relative simplicity to consist in is that such creatures don't seem to weigh reasons in this way at all. Thus, though it is plausible that simpler creatures avoid sensory pains, it is unclear that they 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

²⁹That we have left these questions unanswered would be a shortcoming in what we say if, like the suggestions canvassed in §3, we were attempting to explain away the disconnect reductively in terms of other phenomena.

To be clear, we believe there are answers to such questions — indeed, we believe that any adequate account of suffering will ultimately have to provide some account of such features (Aydede and Fulkerson 2014). But since we don't need to settle these issues to secure what we are claiming is a datum — viz., that there is a disconnect between states of suffering (however construed) and our broader rational lives, there's no reason for us to take sides in these controversies here. (Cf. Aydede and Fulkerson (2015) for an attempt to build an adequate metaphysics of pain from the dual motivational structure we note here.)

that they deemed not in the service of some overall good reason.³⁰ 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: they only have the experience-directed reasons. 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.³¹

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 fully integrated rational control. Such creatures would not need experiential states in order to act in appropriate ways: they would, unlike us, be strongly motivated solely in virtue of the broader practical reasons. In addition, we can imagine that such creatures 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, since (by hypothesis) there would be no motivational remainder. There would just be the informational signal that something bad was going on (and this signal would be appropriately discounted when that information was deemed incorrect). Plausibly, suffering would be motivational in these hyper-rational creatures in just the way (whatever that is) that non-suffering reasons can be motivational to us in typical cases. 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 systems for responding to suffering present in simpler creatures and the flexible, rational system for responding to reasons present in the imagined hyper-rational creatures, and because these distinct

³⁰These 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.

³¹That said, there could be a distinct kind of conflict between reactions to forms of primary 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 motivational role of suffering is exclusively experience-directed. 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.

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 your suffering lose its motivational force.

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 the motivational force of their suffering.

4.3 Taxonomy

In the absence of an uncontroversial account of what suffering amounts to, we have had no choice in the foregoing but to constrain our story about the disconnect between suffering and rationality by appeal to paradigm cases. That is, we have aimed for an account that predicts a disconnect in what we take to be paradigm instances of suffering (e.g., Vaccination) and that predicts no disconnect in what we take to be paradigm non-instances of suffering (e.g., Beach). However, as Rawls (1971, I.4, I.9) famously observed, theoretical equilibrium works in the other direction as well: general principles designed to accommodate cases can themselves suggest revisions to the intuitive taxonomy of cases, particularly once we move away from the most central paradigms. We believe that what we have said about the disconnect between suffering and rationality has a number of taxonomic consequences that fall under this general umbrella.

First, and as we've emphasized throughout, the story we're telling is designed to extend uniformly to the many disparate forms of sensory and non-sensory suffering — including (for example) pain, shyness, hunger, thirst, and certain pathologies of deliberation. The commonality potentially revealed by so extending our story suggests that there may be principled theoretical grounds for counting all of these as instances of suffering, and indeed that there is an interesting theoretical continuity behind what might otherwise seem a dauntingly heterogeneous category (cf. the unity worry mentioned in fn. 4).

Second, as we argued in §2, the very same story extends to chronic pains and psychological compulsions (and underwrites needed explanations of some of the most significant negative consequences of these disorders). And this suggests that

there is something explanatorily useful (besides mere continuity) to be gained from construing the latter as instances of suffering.

A further consequence of what we have said that may initially seem less appealing is that it threatens to extend the category of suffering to even very minor annoyances.³² For example, it seems that our disconnect with rationality could arise even for my mild state of annoyance on hearing elevator music — that state might give me a reason to direct my attention to something (anything) other than the music, but does not diminish in the face of whatever reasons I bring to bear (telling myself that the annoyance is very slight, that it will soon pass, etc.). Still, despite its patterning with our other cases, it might seem overblown to describe this minor annoyance as an instance of suffering. Our response to this consequence is to accept it. In our view, the cost of revising our intuitive taxonomy of suffering to include minor annoyances is outweighed by the view's explanatory benefits. And in any case, nothing we have said precludes acknowledging differences in intensity, duration, import (etc.) among instances of suffering, or, therefore, from distinguishing minor from major instances of the category.

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 our broader 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, and in multiple ways. 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, and between two different kinds of reason, 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.³³

³²Thanks to Michael Rescorla for raising this issue.

³³We are grateful to Marilyn McCord Adams, Craig Agule, David Bain, Brock Bastian, Michael Brady, David Brink, Jennifer Corns, Matthew Hanser, Ross Hetherington, Tom Johnstone, Fiona Macpherson, Alisa Madrigal, Manolo Martinez, Dana Nelkin, Michael Rescorla, Abraham Sapién Córdoba, Assaf Weksler, and Aaron Zimmerman for discussion and comments on earlier versions of this material. We have also benefited from

References

- Abramson, L. Y., Seligman, M. E., and Teasdale, J. D. (1978). Learned helplessness in humans: critique and reformulation. *Journal of abnormal psychology*, **87**(1), 49.
- Amit, Z. and Galina, Z. H. (1986). Stress-induced analgesia: adaptive pain suppression. *Physiological Reviews*, **66**(4), 1091–1120.
- Anscombe, G. E. M. (1957). *Intention*. Harvard University Press.
- Auvray, M., Myin, E., and Spence, C. (2010). The sensory-discriminative and affective-motivational aspects of pain. *Neuroscience & Biobehavioral Reviews*, **34**(2), 214–223.
- Aydede, M. (2014). How to Unify Theories of Sensory Pleasure: An Adverbialist Proposal. *Review of Philosophy and Psychology*, **5**(1), 119–133.
- Aydede, M. and Fulkerson, M. (2014). Affect: Representationalists' headache. *Philosophical Studies*, **170**(2), 175–198.
- Aydede, M. and Fulkerson, M. (2015). Reasons and theories of sensory affect. In D. Bain, M. Brady, and J. Corns, editors, *The Nature of Pain*.
- Bain, D. (2013). What makes pains unpleasant? *Philosophical Studies*, **166**(1), 69–89.
- Ballard, K. and Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage*, **45**(1), 143–150.
- Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, **50**(1), 7–15.
- Bechara, A., Damasio, H., Damasio, A. R., and Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience*, **19**(13), 5473–5481.
- Benedetti, F., Amanzio, M., Vighetti, S., and Asteggiano, G. (2006). The Biochemical and Neuroendocrine Bases of the Hyperalgesic Nocebo Effect. *Journal of Neuroscience*, **26**(46), 12014–12022.
- Benedetti, F., Carlino, E., and Pollo, A. (2010). How Placebos Change the Patient's Brain. *Neuropsychopharmacology*, **36**(1), 339–354.

discussions of this material in presentations at the University of Glasgow and the University of California, Santa Barbara, and on the Value of Suffering blog (<http://www.valueofsuffering.co.uk/is-compulsion-a-form-of-suffering/>).

- Brand, P. and Yancey, P. (1993). *Pain: the Gift Nobody Wants*. Harpercollins.
- Clark, A. (2005). Painfulness is not a quale. In M. Aydede, editor, *Pain: New Essays on its Nature and the Methodology of its Study*. Cambridge Ma: Bradford Book/Mit Press.
- Cohen, J. and Fulkerson, M. (2013). Affect, rationalization, and motivation. *The Review of Philosophy and Psychology*, **4**(4), 1–16.
- Craig, A. D. and Bushnell, M. C. (1994). The thermal grill illusion: unmasking the burn of cold pain. *Science*, **265**, 252–255.
- Cutter, B. and Tye, M. (2011). Tracking representationalism and the painfulness of pain. *Philosophical Issues*, **21**(1), 90–109.
- Damasio, A. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. Avon, New York.
- Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, **60**(23), 685–700.
- Davidson, K., Norrie, J., Tyrer, P., Gumley, A., Tata, P., Murray, H., and Palmer, S. (2006). The Effectiveness of Cognitive Behavior Therapy for Borderline Personality Disorder: Results from the Borderline Personality Disorder Study of Cognitive Therapy (BOSCOT) Trial. *Journal of Personality Disorders*, **20**(5), 450–465.
- Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. MIT Press, Cambridge, Massachusetts.
- Feldman, F. (2002). The Good Life: A defense of attitudinal hedonism. *Philosophy and Phenomenological Research*, **65**(3), 604–628.
- Fulkerson, M. (2014). *The First Sense: A Philosophical Study of Human Touch*. MIT Press, Cambridge, Massachusetts.
- Gendler, T. S. (2008a). Alief and belief. *Journal of Philosophy*, **105**(10), 634–663.
- Gendler, T. S. (2008b). Alief in action (and reaction). *Mind and Language*, **23**(5), 552–585.
- Goldman, A. I. (1970). *A Theory of Human Action*. Princeton University Press.
- Hadjistavropoulos, T. E. and Craig, K. D., editors (2004). *Pain: Psychological Perspectives*. Psychology Press.
- Hall, R. J. (2008). If it itches, scratch! *Australasian Journal of Philosophy*, **86**(4), 525–535.

- Hardcastle, V. G. (1999). *The Myth of Pain*. MIT, Cambridge, Massachusetts.
- Heathwood, C. (2006). The reduction of sensory pleasure to desire. *Philosophical Studies*, **133**(1), 23–44.
- Helm, B. W. (2002). Felt evaluations: A theory of pleasure and pain. *American Philosophical Quarterly*, **39**(1), 13–30.
- Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., and Fang, A. (2012). The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. *Cognitive Therapy and Research*, **36**(5), 427–440.
- Holton, R. (1999). Intention and weakness of will. *The Journal of Philosophy*, **96**(5), 241–262.
- Klein, C. (2007). An imperative theory of pain. *The Journal of Philosophy*, **104**(10), 517–532.
- MacKillop, J., Amlung, M. T., Few, L. R., Ray, L. A., Sweet, L. H., and Munafò, M. R. (2011). Delayed reward discounting and addictive behavior: a meta-analysis. *Psychopharmacology*, **216**(3), 305–321.
- Martínez, M. (2011). Imperative content and the painfulness of pain. *Phenomenology and the Cognitive Sciences*, **10**(1), 67–90.
- McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, **306**(5695), 503–507.
- McKay, R. T. and Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, **32**(6), 493.
- Ossipov, M. H., Dussor, G. O., and Porreca, F. (2010). Central modulation of pain. *Journal of Clinical Investigation*, **120**(11), 3779–3787.
- O’Sullivan, B. and Schroer, R. (2012). Painful reasons: Representationalism as a theory of pain. *Philosophical Quarterly*, **62**(249), 737–758.
- Rawls, J. (1971). *A Theory of Justice*. Harvard University Press, Cambridge.
- Schroeder, T. (2005). Moral responsibility and tourette syndrome. *Philosophy and Phenomenological Research*, **71**(1), 106–123.
- Searle, J. (1983). *Intentionality*. Cambridge University Press, Cambridge.

- Sellars, W. (1956). Empiricism and the philosophy of mind. In H. Feigl and M. Scriven, editors, *Minnesota Studies in the Philosophy of Science*, volume 1, pages 253–329. University of Minnesota Press, Minneapolis.
- Setiya, K. (2007). *Reasons Without Rationalism*. Princeton University Press.
- Silvestrini, N. and Rainville, P. (2013). After-effects of cognitive control on pain. *European Journal of Pain*, **17**(8), 1225–1233.
- Smith, M. (1994). *The Moral Problem*. Blackwell, Oxford.
- Smuts, A. (2010). The feels good theory of pleasure. *Philosophical Studies*, **155**(2), 241–265.
- Teuscher, U. and Mitchell, S. H. (2011). Relation between time perspective and delay discounting: a literature review. *The Psychological Record*, **61**, 613–632.
- Thunberg, T. (1896). Förmimmelserne vid till samma ställe lokaliserad, samtidigt pågående köld-och värmeretning. *Uppsala Läkfören Förh.*, **2**(1), 489–495.
- Velleman, J. D. (1992). What happens when someone acts? *Mind*, **101**(403), 461–481.
- Young Casey, C., Greenberg, M. A., Nicassio, P. M., Harpin, R. E., and Hubbard, D. (2008). Transition from acute to chronic pain and disability: A model including cognitive, affective, and trauma factors. *Pain*, **134**(1-2), 69–79.