

A New Problem for Hard Incompatibilism

Abstract: Following P. F. Strawson (1962), a number of philosophers have adopted and defended a thesis that concerns the connection between hard incompatibilism or, as it's more commonly called, "free will skepticism," and our personal relationships. According to the Strawsonian Thesis, if hard incompatibilism turns out to be true, then its truth would undermine the justification or value of our relationships with other persons. In this paper, I defend the Strawsonian Thesis. Instead of focusing on responsibility-entailing emotions like previous defenses of the Strawsonian Thesis, I argue that if hard incompatibilism is true, we cannot make sense of: (i) the possibility of promissory obligation, (ii) the normative significance of consent, or (iii) the pro tanto wrongness of paternalistic intervention. Because these practices and normative commitments are central to the structure and maintenance of our relationships as we currently conceive of them, it follows that hard incompatibilism has radically revisionary conclusions. By itself this does not mean that no form of free will skepticism is true, but it does suggest that some standard ways of framing the view rest on a mistake.

1. A Strawsonian Thesis

Following P. F. Strawson (1962/2003),¹ a fair number of philosophers have adopted and defended a thesis that concerns the connection between *hard incompatibilism* or, as it's more commonly called, "free will skepticism," and our personal relationships.² According to the *Strawsonian Thesis* (as I will call it), if hard incompatibilism turns out to be true, then its truth would adversely affect our relationships with other persons.

Of course, just what "adversely affect" comes to differs among those who accept the Strawsonian Thesis. So, for example, Susan Wolf tells us that if humans lack free will, "it is obvious why the words "friendship" and "love" applied to relationships in which admiration, respect, and gratitude have no part, might be said to take on a hollow ring," (Wolf 1981, 391). Similarly, Laura Ekstrom claims that if "human beings are wholly without free will (of the sort required for moral responsibility) ... we [should] give up some of the satisfaction that we derive from our relationships," (Ekstrom 2000, 12). And more recently, Seth Shabo has also endorsed a version of the Strawsonian Thesis, claiming that, "our involvement in interpersonal relationships ensures our continued susceptibility to the [responsibility-entailing] reactive attitudes," (Shabo 2012, 97). No doubt, "hollow rings,"

¹ Page references throughout the text will refer to Strawson (2003).

² Cf. Wolf (1981), Anglin (1990), Kane (1996), Ekstrom (2000), and Shabo (2012).

² Cf. Wolf (1981), Anglin (1990), Kane (1996), Ekstrom (2000), and Shabo (2012).

“satisfaction,” and “continued susceptibility” are not the same things, but they each point in the same general direction.

But what direction is this exactly? Well, first it’s important to note that the Strawsonian Thesis doesn’t claim that hard incompatibilism renders it literally impossible for us to engage in personal relationships of the sort that are quite familiar—friendships, collegial relationships, romantic relationships, and so on. No one—not Strawson, nor any of the defenders of the Strawsonian Thesis—think that we couldn’t have such relationships given the truth of hard incompatibilism. Rather, they each seem to mean that something about the *value* or *justification* of those relationships would be undermined, since, for example, relating to another individual as a friend *seems* to require that you first relate to her as a free and responsible agent.

Yet even this weaker statement of the Strawsonian Thesis is not without its detractors. Most notably, Derk Pereboom (2001, 2014) has argued against several recent attempts to defend this claim (including the ones I’ve referenced above).³ In short, the debate between Strawsonians and those who deny the link between responsibility and interpersonal relations can be summarized as follows. Those who accept the Strawsonian Thesis think that if hard incompatibilism is true, then participating in the rich network of human relationships that give meaning to our lives is no longer warranted. Strawsonians thus claim that truth of hard incompatibilism has radically revisionary entailments concerning the legitimacy and value of our relationships with others.

By contrast, hard incompatibilists like Pereboom reject this, arguing that the justificatory grounds of human relationships do not depend on morally responsible agency. Indeed, the world that Pereboom envisions is not the cold, nearly unthinkable world of the so-called “objective attitude,” i.e., a perspective of estrangement and detachment that is putatively opposed to human feelings and relations. Nor is it a world that is devoid of kindness, empathy, and compassion. Moreover, Pereboom tells us, other people are not to be manipulated to suit our needs; their personhood is not denied. Instead, we can be emotionally engaged with them, and in some cases, our lives can be deeply intertwined.⁴

The truth of hard incompatibilism does entail, of course, that we must no longer regard others as *apt* targets of responsibility-entailing emotions like resentment or

³ In addition to Pereboom, see Sommers (2007, 2012), Milam (2016), and Smuts (2016).

⁴ For more on this point, see especially Sommers (2007).

indignation, and that we must also recognize that they are not truly meritorious for their good actions (this is just what the thesis of hard incompatibilism comes to, after all). And this, perhaps, is a loss. But it's not obvious that this is so, especially in the case of emotions like resentment and indignation, which just as frequently poison our relationships with others as they do restore them. If the Strawsonian Thesis is true, then its defenders must have something to say about the plausible alternative offered by hard incompatibilists.

In what follows, I defend the Strawsonian Thesis. However, instead of focusing on responsibility-entailing emotions like previous defenses of the Strawsonian Thesis, I will argue that if hard incompatibilism is true, we cannot make sense of: (i) the possibility of promissory obligation, (ii) the normative significance of consent, or (iii) the pro tanto wrongness of paternalistic intervention. Because these practices and normative commitments are central to the structure and maintenance of our relationships as we currently conceive of them, it follows that hard incompatibilism has radically revisionary conclusions. By itself this does not mean that no form of free will skepticism is true, but it does suggest that some standard ways of framing the view rest on a mistake. Before discussing this new set of Strawsonian-inspired problems for hard incompatibilism, though, I want to first consider the traditional defenses of the Strawsonian Thesis and (following Pereboom) explore their limitations.

2. Relationships and Reactive Emotions

The Strawsonian Thesis has, to many, seemed easiest to defend in the case of love relationships between adults (here I'm thinking of close friendships and at least some forms of committed romantic relationships).⁵ There are several ways that one could develop this claim. But most extant defenses of the Strawsonian Thesis, at least as it concerns relationships of this sort, rely on the fact that the kind of love that exists between partners and friends necessarily has two properties that seem at odds with hard incompatibilism. First, these relationships are *personal*. And second, they have *reciprocity* as a normative ideal.

Concerning this first property, Seth Shabo (2012) has plausibly argued that to love someone, you must take their actions personally. This means, among other things, that you must *care* about what they do, and in particular, that you must care about how their actions

⁵ For a longer defense of this, see Shabo (2012).

affect you. But if you take their actions personally in this way, then, as Shabo argues, you'll naturally and invariably be disposed towards resentment in cases in which their actions manifest insufficient regard. Love, it seems, must involve ineliminable psychological dispositions for resentment and other responsibility-entailing emotions.

There is, I think, something to this line of argument. But as Pereboom argues, it is ultimately deficient. Citing relationships with young children as but one example, Pereboom notes that, in fact, there are actually many "relationships very important to us in which our care for others' attitudes is personal, but in which many of us are not subject to resentment," (Pereboom 2014, 185). And this seems right. For example, my feelings are hurt when my young daughter pushes me away before bedtime. This reflects the fact I care deeply about her attitudes towards me. But despite this deep concern, I have to say, I'm never tempted by resentment.⁶ I take her actions personally, and yet, I am reliably able to forgo resentment.

So too, in my dealings with students, I take their actions personally. I care deeply about how they treat me and their other classmates, and also, I care about how seriously they regard the material that we're discussing together. But again, as was the case with my daughter, I am not particularly prone to resent them when they act callously or lazily or foolishly in some other way. They are still young and figuring things out, after all. Furthermore, it's the first time many of them have ever had much freedom to make decisions (and consequently, mistakes) on their own. And, truth be told, I also care about them deeply as people, and this makes me something of a "softie." So although I do take my students and their actions personally, rarely—perhaps verging on "never"—do I feel pulled towards real anger of any sort.

This means that the Strawsonian cannot defend the legitimacy of the reactive attitudes by claiming them to be *psychologically* necessary for close personal relationships. Of course, here the Strawsonian will be quick to observe that the personal relationships you might have with a young child or with students that you care about are importantly different than the relationships of friendship or of "the sort of love which two adults can sometimes be said to feel reciprocally, for each other," (Strawson 2003, 79). These latter forms of human involvement aren't just *personal*, like our relationships with young children or students; they also involve a kind of mutuality or reciprocity of expectations that each party

⁶ Of course, my daughter is only two right now, so this will probably change. Perhaps, then, the claim that we are unavoidably susceptible to resentment because we take others' actions personally might be true in the case of our relationships with adults.

to the relation must hold the other to. So although the hard incompatibilist is right to reject defenses of the Strawsonian Thesis that rest solely on the fact that our love relationships are personal, perhaps the Strawsonian Thesis can be defended by appeal to a further property of the set of relationships in question: the fact that they require that the participants hold one another to norms of reciprocated respect.

From the reciprocal nature of friendships and of the mature love relationships adults sometimes find themselves in, one can straightforwardly (but as we'll see, not *uncontroversially*) offer a defense of the Strawsonian Thesis. After all, if each of the parties to such relationships must hold themselves and the other to norms of reciprocated respect, then each party must be prepared to hold themselves or the other responsible for violations of this norm. But holding others responsible is (very plausibly) constitutively linked to the desert-entailing reactive emotions of resentment, indignation, and guilt.⁷ So part of what it is to love is to be prepared to resent your beloved in cases in which he or she has acted poorly. However, if hard incompatibilism is true, then I have reason to eschew such preparations. Accordingly, if hard incompatibilism is true, then I have reason to eschew the very relationships that require I be prepared to hold the other party morally responsible.

Here again, Pereboom demurs. Specifically, he argues that reciprocal love relationships that eschew putatively desert-entailing emotions like resentment or indignation aren't obviously flawed. The reason for this, on Pereboom's view, is that although it matters for a reciprocal relationship that each party is committed to holding the other to certain norms and expectations, it doesn't follow that this can only be done through desert-entailing reactive emotions of the sort that are illicit if hard incompatibilism is true. So even if some forms of *holding responsible* cannot be pruned from our reciprocal love relationships, why *must* love require us to hold others responsible through basic desert-entailing emotions like resentment or indignation?

And despite the insistence of Strawson and his followers (Wallace 1994; Helm 2011; author citation), it is plausible that one can hold another to the norms of reciprocated respect without a proneness to resentment or indignation. I can, for instance, hold you responsible for cheating at a game simply by penalizing you. This need not indicate or express resentment in any way. More generally, as Pereboom (2013, 2014) argues, there are

⁷ See Wallace (1994, 2011) for a full defense of this claim.

forms of blame and interpersonal moral address not rooted in resentment or indignation that can nevertheless serve to hold others responsible for their actions.

Developing this point, he says:

The free will skeptic would resist or disavow resentment and indignation, but she would not be exempt from disappointment, sorrow, and concern for the offender upon being wronged. She would have remorse for her own immoral actions grounded in sympathy with those affected, and the moral resolve she would take effective measures to eliminate dispositions to such actions, reconcile with those she has wronged, and to restore her relationships. When hurt by another, she might blame in the forward-looking sense I set out [in Pereboom 2013, 2014], but upon a commitment on the part of the other to eliminating the disposition to act this way, she would acknowledge this commitment and cease to regard the hurt as an obstacle to her relationship, (Pereboom 2014, 192-93).

The skeptic, as Pereboom sees him, can be fully, or almost fully engaged, with others reciprocally without being prone to experience responsibility-entailing reactive emotions. Neither these emotions nor moral responsibility (in the basic desert-entailing sense) appear to be conditions for reciprocity.

Whether this reply succeeds as an answer to the Strawsonian Thesis turns on how we conceive of our interpersonal relationships. I think that one thing Strawsonians might be taking for granted is that there is a kind of direct and unmediated emotional engagement that's central to reciprocal relationships—particularly reciprocal love relationships between adults. (Strawson's talk of the "non-detached" reactive attitudes certainly invites this reading.) To sever this connection on the basis of worries of the sort that motivate hard incompatibilism—worries that Strawson would no doubt take to be mere metaphysical flights of fancy—brings not only one but several thoughts too many into our relationships. The hard incompatibilist sketched above might be able to participate in reciprocal relationships, but given the alienation that seems to exist between natural human engagement and his metaphysical commitments, it's hard to see how those relationships can really be meaningful in the way we characteristically take them to be.

To this, the hard incompatibilist can rightly say that there is a great deal of “distance” and calculation that already goes into our personal relationships. In fact, the very idea that there is some unmediated emotional connection between persons is overly sentimental. How we actually relate to one another is a great deal more complicated than this picture of human relations suggests—a fact that the breadth of human emotional experience already attests to. Moreover, to flatly deny that the picture Pereboom sketches above can involve *meaningful* reciprocal relationships seems problematically conservative. The emphasis on responsibility-entailing emotions like resentment or indignation is a relatively recent Western preoccupation. But surely we’re not the only ones who have found our personal relationships to be meaningful. Perhaps there are radically different forms of reciprocal engagement that are also good and meaningful.

Here, it seems, the arguments (begin to) run out. The Strawsonian will again simply want to insist that we cannot hold others to norms unless we’re prepared to resent them (or feel indignation towards them) if they violate that norm. Imagine, for example, someone who *wasn’t even disposed* to resent others when she was insulted. Is it really credible that she holds them to the norm that they don’t do so? Her emotional equanimity suggests that she doesn’t, that what happens doesn’t really *matter* to her. But if how others treat her doesn’t matter to this person, then how can her relationships with those individuals be a source of meaning in her life?

Of course, hard incompatibilists like Pereboom will again want to insist that this is overblown. We can and do hold others responsible for norm violations without resentment or indignation, or even the disposition to feel those emotions. Gandhi and Martin Luther King, Jr., for example, were very much in the thick of things—no one could describe either of these men as failing to hold others responsible—and yet they eschewed rhetoric of the sort that might incite resentment or unbridled anger.⁸ So, we too, can hold others responsible without responsibility-entailing emotions like resentment or indignation. The dividing line between these two views is clear, even if the path to resolution isn’t apparent.

What this means, I think, is that we’ve reached what John Martin Fischer has called “a dialectical stalemate.”⁹ That is, we’re at the point in a debate where we’ve reached a kind of intuitive standoff between Strawsonians and hard incompatibilists. The former thinks that

⁸ See Watson (1987) for more discussion on this point.

⁹ Fischer (1994).

(i) reciprocity requires a willingness to hold responsible and (ii) a willingness to hold responsible necessarily involves responsibility-entailing emotions. The latter denies this last bit. Here we have two distinct *pictures* of what it means to hold someone responsible for a norm violation. And plausibly, no argument could offer decisive grounds for preferring one picture to the other. (Or weaker: no argument should be held to this standard.) Let me quickly say that this isn't a *failure* on the part of either Strawsonians or hard incompatibilists; it's just a fact of discursive forms of philosophical conversation.

Summing up these points (in a slightly different domain), Fischer rightly says, "wisdom in philosophy consists partly in recognizing that one should not expect decisive arguments in most contexts. Rather, strong plausibility arguments are all that it is reasonable to hope for," (Fischer 1994, 85). This is sage advice. So for the remainder of this paper, I want to simply table issues that concern responsibility-entailing emotions and their relationships to reciprocal interpersonal relationships, since it's clear that no decisive arguments will be on the offing in this debate. Instead, I will focus to a set of social practices and normative commitments that are also bound up in reciprocal interpersonal relationships, with the aim of doing what Fischer suggests, viz., offering a strong plausibility argument in defense of the Strawsonian Thesis.

3. Hard Incompatibilism and Social Practice

Even if we set aside issues concerning responsibility-entailing emotions, more needs to be said in order to vindicate Pereboom's conclusion that hard incompatibilism poses no threat to normative bases of our personal relationships. After all, there are a great many social practices and institutions that affect the meaning and value of these relationships. Indeed, we understand and value these relationships not in isolation, but as part of a larger web of human involvement. But if the normative commitments that underwrite those social practices and institutions are themselves inconsistent with hard incompatibilism, then it will follow that skepticism about moral responsibility does in fact have significant implications for the meaningfulness of our personal relationships.

With that in mind, I now want to turn to three things that apparently require morally responsible agency and that would be ruled out on the truth of hard incompatibilism: the possibility of promissory obligation, the normative significance consent, and the pro tanto wrongness of paternalistic intervention.

3.1. Promissory Obligation

Suppose I promise you that I'll help you move but I then fail to deliver. Now suppose that you discover that I made that promise only under extreme duress. Plausibly, you would conclude that my utterance "I promise that ..." was infelicitous in some crucial way and that therefore, it wasn't binding. That is, if I promised only as a result of extreme duress, then I couldn't come to be obligated to keep that promise even though I (apparently) performed a speech act that, *in other conditions*, would be sufficient to create an obligation.

Similarly, if you find out that I've been coerced to make the promise, or if I only made the promise because I was forced to take a drug that made me especially susceptible to agree to whatever others asked of me, or ... you'd probably conclude that I was not morally obligated to keep the promise. So too, if I was a young child, who was only at the very earliest stages of understanding the social and normative significance of "I promise."

In all these cases, I think there's a simple and unified explanation for why the outward utterance, "I promise to help you move" fails to actually obligate: none of the agents' promises were made of their own free will. Unlike categorical moral obligations, which (putatively) bind us simply because we are moral agents, promissory obligations are voluntarily undertaken. This means that the promisor's will has to be implicated in her promise *in the right way*. And when the promisor is under extreme duress, or coerced, or ..., then her will is not free, and she does not come to be obligated to follow through on the promise that she apparently makes. But notice: the hard incompatibilist tells us that, although normal adult humans are not identical to those under duress, subject to coercion, or ..., their actions are similarly unfree. After all, on hard incompatibilism, we are no more the source of our actions than any of the characters assembled above. If hard incompatibilism is true, it seems true of no one that they have made a promise "of their own free will." However, this is just to insist that no one has ever genuinely been subject to a promissory obligation. Hard incompatibilism is apparently at odds with the very idea of promissory obligation.¹⁰

It is no coincidence that the very sorts of impairments that ordinarily serve as excusing or exempting conditions on an agent's status as free in the sense required for

¹⁰ I say "apparently" here because, as I'll discuss in §5, the hard incompatibilist will no doubt want to object to this argument.

morally responsible (i.e., duress, coercion, childhood, etc.) *also* serve to sever the link between the speech act of promising and the creation of a promissory obligation. The possibility of promissory obligation and an agent's being morally responsible for some action must each require that the agent's will be implicated in the promise/action in just the same way. And though this point has been largely ignored, it seems just right. Consider, for example, what it would mean to deny this explanation. It would mean that, say, our agency is significant enough to (in principle) genuinely obligate us to perform incredibly difficult, demanding, and burdensome tasks, and yet it is not significant enough to render us deserving of attitudes like resentment or indignation should we fail to meet our obligations. This is puzzling.

This is all rather unfortunate for hard incompatibilists like Pereboom, who want to maintain that their skepticism about moral responsibility is a minimally revisionist position, since promises play an important role in our lives. Informal promises of the sort that we make on the playground as children are perhaps not that important—though don't tell this to the person who's promised not to tell your classmates who you have a crush on. But the promises we make as adults, to our partners for example, are quite significant. They are *inter alia* expressions of our love. But unlike other expressions of our love (like "Jimmy loves Sarah" painted on the town water tower), the promises themselves constitute a deep form of commitment, since they, unlike other expressions of love, *obligate* us to continue in our love.

And it's not just promises that we have to worry about. Contractual obligation serves as a formal, institutional analog to promissory obligation since its moral justification lies, in part, in the idea that promissory obligation is a special and significant form of obligation. But just a contract that is coercively executed cannot be enforced legitimately, it's generally difficult to see why any contract that is executed by an agent who is not free or responsible can be enforced. No doubt, one could try to justify contractual obligation and contract law more generally in purely consequentialist considerations—it's good, after all, to have institutions that enable people to make and enforce agreements of a certain kind. In such a case, perhaps strict liability can serve as the standard of enforceable contracts, and hard incompatibilism is consistent with that standard.

The problem with this, however, is that although a framework for contractual law that operates with the standard of strict liability for contract enforcements might be a good institution for a state to establish, it's not obviously a *fair* institution, since it's not clear that I

should be responsible for bearing the burdens of a contractual obligation if I did not freely enter into that obligation. Further, if one admits that the *only* rationale for enforcing contracts is consequentialist, then one has no principled reason not to intervene in contractual situations in which doing so will greatly benefit one party without significantly harming the other. But this is just to deny that there is some special moral category of contractual obligation.

The hard incompatibilist is, of course, free to deny this, but such denials surely come at a cost, since one must then be prepared to revise large portions of tort law and actual social practice. (Or else, find a different justification for those portions of tort law.) If hard incompatibilism upsets the possibility of promissory (and so, contractual) obligation, then clearly, it upsets our personal relationships in a rather significant way.

3.2 Consent

You might think that this overstates things, since perhaps we make promises only quite rarely.¹¹ But the problem that emerges above generalizes in worrisome ways. Consider now the role that consent and agreement play in our personal relationships. Part of what makes our relationships *reciprocal* is that each party's consent and agreement is important. Though we can do stuff together, we can't be *friends* if I never take your consent (or lack thereof) to be normatively significant. But hard incompatibilism is at odds with thinking that consent is significant in this way.

To see this, let's suppose I request that you let me live with you rent-free for a month while I look for a new place to live. Ordinarily, it's wrong to just show up at someone's house and squat there for a month, so I can't move in solely of my own accord. However, if I secure your consent, then I've done nothing wrong by living there, since you've temporarily granted me that right. But now suppose that my parents, who don't want me to move back in with them, were extorting you. They tell you that if you don't let me stay at your place for a month, they are going to break all of your fingers and toes (and if that doesn't work, they say they're going to move on to other joints). In that case, the mere fact that you agreed to me moving in doesn't seem morally significant. If your consent had been freely given, then it'd be permissible for me to live in your house, but because it wasn't freely

¹¹ Or never at all if Elinor Mason (2005) is correct.

given, it doesn't seem permissible. (Of course, if I didn't know about my parents' intervention, perhaps I'd be excused for living there.)

Or instead suppose that I secured your agreement by getting you really drunk, knowing that you're exceptionally generous while intoxicated. Again we might imagine that you say, "yes" in response to my request, but again it seems that I have failed to receive a morally significant form of consent, since mere verbal agreement is insufficient for that. And though hard incompatibilism isn't the thesis that ordinary adults are no different than intoxicated agents or the thesis that ordinary adults are no different than agents who are being extorted, it does entail that none of us have free will.

Now it's true that Pereboom himself focuses on the connections between lacking free will and deserving praise or blame, but as these cases suggest, lacking free will has further implications for our relationships with others. That I am not free means that my consent doesn't change the normative status of your actions in the ways we typically take it to, since what really matters is not just consent, but *freely given* consent. But if hard incompatibilism is true, then no consent is freely given. Consequently, if hard incompatibilism is true, in assessing whether I should move in with you, what matters is not your will, but rather the overall goodness of that state of affairs. After all, if my moving in is a Pareto optimal situation, then, if we suppose that consent lacks the normative significance that we normally attach to it, it's hard to see what reason I have *not* to move in if that's what I want to do.

This picture of the normative significance of consent has its roots in Kantian (or perhaps contractualist) accounts of morality. The idea that we should never treat others as mere means, but always as ends in themselves is one that puts a lot of weight on the importance of consent. After all, what makes it okay that I tell the waiter to get me a refill—i.e., what makes it true that I'm not using him as a *mere* means—is that he has freely consented to the conditions of his employment.¹² This suggests an even more worrisome point, viz., that hard incompatibilism is in deep, perhaps irreconcilable tension with Kantian-inspired moral theories. If it follows from hard incompatibilism that actual consent lacks moral significance, then we have reason to reject moral theories that simply presuppose or are otherwise committed to its centrality to the moral life.

¹² Of course, here we must also imagine (per impossible?) that his consent is not made under the kind of duress that Marxists imagine to be lurking in the background of all post-industrial capitalist societies.

In response to this way of framing the worry, Pereboom argues that the kind of agency that Kant requires for moral agency is importantly distinct from the kind of agency that is required for moral responsibility.¹³ The former form of agency requires that we be capable of “setting ends and choosing means” and for “formulating [practical] principles and making commitments to them,” (Pereboom 2001, 151). The latter form of agency, the hard incompatibilist insists, requires that we be ultimate causal sources of our actions. But even if no one is the ultimate causal source of their actions, it won’t follow that no one is an ends-setter or capable of normative commitment. “The capacities for these activities,” Pereboom says, “can remain intact [even if hard incompatibilism is true],” (Pereboom 2001, 151).

But this, I submit, does not vindicate the normative significance of consent. The reason for this is related to a point made earlier (in §3.1). Because the excusing and exempting conditions on moral responsibility are apparently isomorphic to the conditions under which consent has diminished significance, it seems that the kind of agency implicated by our ordinary conception of the significance of consent is no different than the kind of agency that underwrites our status as morally responsible agents. And if there is a difference, then the hard incompatibilist owes an account of this difference.

Furthermore, there is a second reason to think that the agential capacities that give us reason to respect others’ consent are no different than the ones that underwrite morally responsible agency. Consent is normatively important because it can affect how I have *reason* to treat you. If you agree to certain kinds of treatment, then it means I can engage with you in potentially burdensome ways. Your agency thus directly licenses otherwise impermissible treatment. But this is precisely what happens when we’re morally responsible for our actions. When you act in, say, a blameworthy fashion, then you give me a reason to treat you in potentially burdensome ways—ways that would otherwise be impermissible. Most notably, you give me reason to resent you. In neither of these cases, however, are the reasons grounded in consequentialist considerations. Rather, the reasons for potentially burdensome treatment/resentment seem directly responsive to facts about your agency. If the hard incompatibilist wants to claim that there are important differences between the agential capacities that underwrite the normative significance of consent and our status as morally responsible agents, then, again, they owe us an account of why reasons to treat someone in

¹³ On this point, see Pereboom (2001), especially pgs. 150-52.

potentially harmful and burdensome ways obtain in virtue of fundamentally different features of our agency than do reasons to blame.

3.3. Paternalistic Intervention

Finally, related points can be made about paternalistic interventions. Typically, we think it's impermissible (or weaker: that there are weighty *pro tanto* reasons not) to intervene in someone else's choices, even when we think those choices are suboptimal, irrational, or just plain stupid. There are, of course, exceptions to this. You can use force to stop your child from putting her hand on the hot stovetop. You can take your friend's car keys and phone when they're drunk—they don't need to drive or send texts to their ex. In the case of some medical interventions, it might (let me emphasize "might") be permissible to ignore the patient's wishes in order to better secure their welfare, at least when the stakes are sufficiently high. But in the case of ordinary interactions with other adult persons, these kinds of interventions—even when they make the person in question better off—seem objectionable.

The explanation of this is simply that if someone's stupid choice is made of their own free will, we take ourselves to have reason to refrain from paternalistic intervention even if such intervention would be good for the individual all things considered. But if hard incompatibilism is true then these powerful reasons to refrain from paternalistic intervention don't arise. This might not entail that we *must* intervene paternalistically when doing so benefits the individual, but it seems to entail that we would have reason to do so—that it would be permissible at least. And even this weaker entailment would have significant implications for our ordinary relationships. If I have reason to engage with other adults in the ways that I should engage with young children or drunk friends, then although I'm not regarding them as objects, I'm certainly not regarding others as people with whom I can have meaningful and mutually nurturing reciprocal relationships. After all, if I regard myself as having reason to paternalistically intervene in your life, even if you have not asked me to do so, then I see myself as (potentially) having reason to manipulate you in order to leave you better off. But insofar as I regard myself as potentially having reason to manipulate you, I cannot really regard myself as being friends with you, since that requires me to view you as an *equal* participant in our relationship.

Worse, it seems that paternalism of this sort is at odds with reciprocal love. Kyla Ebels-Duggan (2008) has recently argued that the practical dimension of reciprocal love—i.e., the dimension of love that accounts for its reasons-giving force—cannot be assimilated to a concern for the beloved’s welfare. Instead, to love someone we must regard them as having “selection authority” over us as well as “authority in judgment,” i.e., we must regard their choices as being a source of reasons for us and their judgment as being *prima facie* valid. In practice this means, e.g., if you love your partner, you must see their desire to go to the opera as a reason to go to the opera, even if it’s the last thing you’d be inclined to do on your own. Of course, you might not end up at the opera. But if you fail to regard yourself as having some reason to be there, then you seem quite unloving in this case. And if this kind of normative indifference characterizes your relationship with your partner more generally, if you never regard their preferences and evaluative judgments as reasons-giving, then it’s implausible to think that you really love them at all.

These features of love explain why it’s typically impermissible (and certainly *inadvisable*) to intervene paternalistically within the context of reciprocal love relationships, even when doing so would make your beloved better off. A commitment to paternalistically intervene whenever it boosts your beloved’s welfare is a commitment to not regard him or her as giving you reasons to aid them in bringing about their ends. In that case, you are prepared to simply ignore the putative authority their choices have over you. But *not* treating their choices in this way is, at least according to Ebels-Duggan, necessary for loving them. So a commitment to paternalistic intervention, which seems warranted on skeptical grounds, is at odds with the practical component of love. Hard incompatibilism might therefore undermine the practical dimension of our reciprocal love relationships, even if we agree with Pereboom that one can have meaningful love relationships without responsibility-entailing emotions.

4. Variations of the Participant Attitude

The relationships that make sense on hard incompatibilism are apparently ones that don’t involve promises, don’t take consent to be normatively significant, and don’t preclude extremely paternalistic forms of intervention and manipulation. I’ve tried to adduce support for this indirectly, but I think it is possible to make this case in a slightly more direct way.

In “Freedom and Resentment,” P.F. Strawson introduces the idea of the “participant attitude,” which is the general attitude we must take towards others in order to participate in interpersonal relationships with them. Strawson himself describes it almost exclusively in terms of the so-called reactive attitudes, which are responsibility-entailing emotions like resentment and indignation (and also: guilt, gratitude, esteem, hurt feelings, pride, etc.). Now, as we saw in §2, hard incompatibilists like Pereboom have developed a powerful challenge to this conception of the participant attitude. If hard incompatibilism is correct, it is at least plausible that the kind of emotional engagement that characterizes human participation and sociality need not involve responsibility-entailing emotions. Indeed, part of why I take Pereboom’s version of hard incompatibilism to be so attractive is that it respects our conception of ourselves as participants in deeply rewarding personal relations.¹⁴

However, by focusing almost exclusively on responsibility-entailing emotions, I think that Strawson (and those following in his wake) have unduly circumscribed the participant attitude. For him, the more general participant attitude seems to be comprised *wholly* of our emotional engagement with others via the reactive attitudes.¹⁵ But if this is right, then it’s very plausible that hard incompatibilism is consistent with, if not the participant attitude as Strawson himself conceived of it, an attitude of interpersonal engagement that involves genuinely personal emotions (albeit not responsibility-entailing ones). Hard incompatibilists like Pereboom would thus be within their rights to conclude that our relationships are, at most, only minimally affected by the truth of hard incompatibilism.

As I see it, however, the participant attitude is a much more general practical orientation. Plausibly, this does involve the rich forms of emotional engagement of the sort that Strawson discusses. Or at least, it involves the forms of emotional engagement that hard incompatibilists like Pereboom and Sommers (2007) describe. But even more broadly, the participant attitude is also a lattice of mental attitudes, deliberative tendencies, and behaviors that structures and organizes our reciprocal relationships with others. That is, it is a much more general framework that facilitates interpersonal engagement with others. This means that even if we grant that the emotional attitudes that figure in this complicated network need not be responsibility-entailing ones, as hard incompatibilists maintain, it won’t follow

¹⁴ This is also true of the skeptical view developed in Sommers (2007).

¹⁵ To be fair, it’s his term, so he can define it however he wants. But insofar as he’s picking a real practical orientation that we have towards others, it seems like characterizing it exclusively in terms of the so-called reactive attitudes is a mistake.

that we can thereby understand other constitutive mental attitudes, deliberative tendencies, and behaviors as not being responsibility-entailing.

For this reason, the hard incompatibilist owes us an account of promissory obligation, the normative significance of consent, and the pro tanto wrongness of paternalistic intervention. Such an account will need to explain why none of these, each of which seems to be part of the lattice that makes up the participant attitude, is responsibility-entailing. Alternatively, the hard incompatibilist owes us an account of why pruning these practices from our lives leaves us with a recognizable and meaningful form of reciprocal social interaction.

Unfortunately for the hard incompatibilist, each of these tasks is difficult. The latter is difficult because it is hard to imagine what is left of our relationships after we prune promissory obligation, deemphasize the importance of consent, and come to regard paternalistic intervention as generally permissible. Though they might be satisfying in some ways and perhaps even a source of comfort, such relationships would resemble friendship and love as we currently understand them in only the barest ways.

On the other hand, the former task will be difficult because, as we've seen, the same conditions that undermine an agent's responsibility (e.g., coercion, duress, youthful naiveté, drunkenness, etc.) also seem to undermine promissory obligation, the normative significance of consent, and the pro tanto impropriety of paternalistic intervention. This means that if there is some general condition that undermines our status as responsible agents, that condition should undermine these other things as well. But insofar as the hard incompatibilist wants to insist that the truth of hard incompatibilism is minimally revisionist, it seems that this is exactly the route they must take. In the remainder of the paper, then, I'll consider one way that the hard incompatibilist might attempt to reply and explain why I think it fails.

5. The Promise of Compatibilism?

In a number of places, Pereboom has employed a powerful strategy for dealing with objections of this general kind. In short, Pereboom claims that when hard incompatibilism is pinned with a seemingly implausible consequence—e.g., that it is incompatible with genuine deliberation or that it can't make sense of holding others to norms (which putatively require a willingness to blame), for example—Pereboom argues that compatibilist accounts of the

phenomena will do the trick.¹⁶ That is, he argues that although compatibilism fails as a theory of moral responsibility, it gives us the resources for fine theories of deliberation and even some forms of blame.¹⁷

I think this strategy really is quite brilliant. And by my lights, it's successful (or very nearly successful) in the cases of deliberation and of holding others to norms by blaming them. Perhaps, then, Pereboom can simply adopt this strategy here, and give us a compatibilist account of promissory obligation, the normative significance of consent, and the pro tanto wrongness of paternalistic intervention. If so, then he can plausibly maintain that the truth of hard incompatibilism doesn't commit him to a strained and impoverished conception of human relationships. I therefore want to conclude by offering an argument against this strategy in these particular cases. For although I think the compatibilist account of deliberation, e.g., is sufficient to account for deliberation as we conceive of it, I'm skeptical that one can endorse hard incompatibilism and a compatibilist account of promissory obligation, the normative significance of consent, or the pro tanto wrongness of paternalistic intervention. Here's why.

6. Manipulating the Four-Case Argument

Hard incompatibilists like Pereboom are sourcehood theorists. According to this view, agents are morally responsible for their actions only if they are the source of those actions. And though there are a variety of routes to being a sourcehood theorist, by far the most common (and most compelling) is via the manipulation argument.¹⁸ Indeed, Pereboom's rejection of compatibilism rests on precisely this sort of argument: the Four-Case manipulation argument. If this argument (or some other variant of the manipulation argument) is sound, then it shows that the conditions that compatibilists take to be sufficient for moral responsibility are not actually sufficient.

But notice: if it's possible to argue against compatibilism (understood as a thesis about the conditions under which agents are morally responsible), it should be possible to run an identical argument against compatibilist accounts of promissory obligation, the normative significance of consent, and the pro tanto wrongness of paternalistic

¹⁶ See Pereboom (2001, 2014), particularly 2014, pgs. 104-152.

¹⁷ Cf. Sher (2006) and Scanlon (2008).

¹⁸ Cf. Wisdom (1934), Pereboom (1995, 2001, 2014), and Kane (1996), among others.

intervention.¹⁹ And if the former argument is sound, as the hard incompatibilist says, then the latter ones must be sound as well. But then, hard incompatibilism would rule out not only desert-entailing moral responsibility but also promissory obligation, et al. If, on the other hand, the latter arguments are not sound (as I think is very clearly the case), then the compatibilist about moral responsibility has good grounds for rejecting Pereboom's original Four-Case argument. The hard incompatibilist thus faces a dilemma: it's either the case that hard incompatibilism requires very unhuman forms of social engagement, or, in an effort to avoid this unpalatable result, it sows the seeds of its own destruction by giving the compatibilist a principled explanation of why the Four-Case argument fails to show compatibilist accounts of moral responsibility to be false.

With this in mind, let me leave you with a schematic version of a modified Four-Case manipulation argument against promissory obligation. For reference, I'll start where Pereboom does. I'll then appropriate the framework that Pereboom gives us for my own purposes. Consider, then, Pereboom's Case 1.

Case 1. A team of neuroscientists has the ability to manipulate Plum's neural states at any time by radio-like technology. In this particular case, they do so by pressing a button just before he begins to reason about his situation, which they know will produce in him a neural state that realizes a strongly egoistic reasoning process, which the neuroscientists know will deterministically result in his decision to kill White. ... his process of deliberation from which the decision results is reasons-responsive; in particular this type of process would have resulted in Plum's refraining from deciding to kill White in certain situations in which his reasons were different. His reasoning is consistent with his character because it is frequently egoistic and sometimes strongly so. Still, it is not in general exclusively egoistic, because he sometimes successfully regulates his behavior by moral reasons, especially when the egoistic reasons are relatively weak. Plum is not constrained to act as he does, for he does not act because of an irresistible desire – the neuroscientists do not induce a desire of this sort.²⁰

¹⁹ Trevor Pisciotta (2009) similarly develops a parody argument against Pereboom's hard incompatibilist account of meaning and fulfillment.

²⁰ Pereboom (2014): 76-77.

Here Plum meets all of the purported compatibilist conditions on moral responsibility, and yet, the hard incompatibilist insists that he is not morally responsible for killing White because that action “is causally determined by the neuroscientists’ intervention, which is beyond his control,” (Pereboom 2014, 77). At this point, it’s not essential that you agree with the hard incompatibilist in their judgment; what’s important is that you appreciate why Pereboom and other hard incompatibilists (reasonably, but erroneously in my opinion) take Plum to be exempt from moral responsibility.

With this in mind, let’s turn to a parallel case in which an agent is manipulated in precisely the way that Plum is manipulated, and consider what the hard incompatibilist must say about that agent.

Case 1*. Mustard is an agent who is just like Plum in Case 1 of Pereboom’s four-case manipulation argument. A team of neuroscientists periodically intervenes in his reasoning processes, but they do so in a way that leaves his rational capacities intact.²¹ Through their interventions, Mustard is causally determined weigh reasons in a way that leads him to promise Scarlet that he will help her move.

The day of the move comes, and we want to know, is Mustard actually obligated to help Scarlet—i.e., would he be doing something objectionable by failing to help? Or does he do nothing wrong if he decides at that time to refrain from helping?

For those of us who know about the neuroscientists’ work, is it plausible that through the speech act of uttering “I promise...” Mustard has come to be obligated to help? I tend to think so, and yet, I can see why someone without a settled view of these matters would be agnostic. But whatever you, me, or an unsettled agnostic might be inclined to say about this case, it’s hard to see how the hard incompatibilist could accept this. For just as it’s tempting to say of Plum that he is exempt from blameworthiness because he is not the source of his action, so too, it seems tempting to say of Mustard that he is not obligated because the promise wasn’t *his* in the relevant sense.

Moreover, if one accepts the soundness of Pereboom’s Four-Case argument, one can also say of Mustard that in this case, he isn’t morally responsible for promising. But can one

²¹ The rational capacities I have in mind here are whatever capacities compatibilists take to be sufficient for moral responsibility.

be obligated to keep a promise that one is not morally responsible for making? I don't see how. As I've already noted, promises are optional but often very burdensome obligations. It doesn't seem fair that someone come to be obligated in this way when you admit that she is not morally responsible for incurring the obligation.

Consider now Case 2*, which parallels Case 2 of Pereboom's Four-Case argument.²²

Case 2*. Mustard is an agent who is just like Plum in Case 2. A team of neuroscientists has programmed him from birth to reason in a particular way. This initial programming secures the neuroscientists' preferred form of reasoning without interfering with Mustard's rational capacities. Given his programming, Mustard is causally determined in the circumstances to weigh reasons in a way that leads him to promise Scarlet that he will help her move.

Again we can ask ourselves, is Mustard obligated in help in Case 2*? Well, as Pereboom points out in his statement of the Four-Case argument, it's hard to see why added temporal distance to the neuroscientists' intervention matters. So if he wasn't obligated in Case 1*, then plausibly, Mustard isn't obligated in Case 2*.

Cases 3* and 4* are just like Cases 3 and 4 of Pereboom's Four-Case argument. They are therefore different than Cases 1* and 2*, but it isn't obvious that their differences are relevant to promissory obligation.²³ And since Case 4* is a case of ordinary causal determination by blind physical forces, it seems that a compatibilist account of promissory obligation won't work for the hard incompatibilism. Just as Pereboom's original Four-Case argument putatively shows compatibilist conditions on moral responsibility to be

²² For reference, here is the text of Pereboom's Case 2.

Case 2. Plum is just like an ordinary human being, except that a team of neuroscientists programmed him at the beginning of his life so that his reasoning is often but not always egoistic (as in Case 1), and at times strongly so, with the intended consequence that in his current circumstances he is causally determined to engage in the egoistic reasons-responsive process of deliberation and to have the set of first and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his actions by moral reasons, but in his circumstances, due to the strongly egoistic nature of his deliberative reasoning, he is causally determined to make his decision to kill. Yet he does not decide as he does because of an irresistible desire. The neural realization of his reasoning process and of his decision is exactly the same as it is in Case 1 (although their causal histories are different), (Pereboom 2014, 77).

²³ Moreover, if there is a relevant difference, then it's hard to see how that same difference can't be exploited by the compatibilist.

insufficient, I think this modification of the four-case argument shows that, for hard incompatibilists at least, compatibilist conditions on promissory obligation are also insufficient.

Notice, however, this argument doesn't pose a problem for the compatibilist, since she either accepts that Mustard is morally responsible for making the promise in Case 1*, and so is not committed to him being obligated even though he is not morally responsible, *or* she takes there to be an important difference between Case 1 and Case 4 of Pereboom's Four-Case argument that allows that Mustard would be morally responsible in Case 4*, which would simply be a case in which Mustard is causally determined to make the promise by blind physical forces. The compatibilist, then, is not committed to the idea that one can be bound to keep a promise that one is not morally responsible for making. The hard incompatibilist, however, is committed to this—at least insofar as he or she wants to maintain that we are sometimes subject to promissory obligation. I find this a tough pill to swallow, and so it seems to me that the hard incompatibilist should instead just accept that we are never subject to promissory obligation. By itself this doesn't mean that hard incompatibilism is false—just that it requires a more revisionary conception of the participant attitude than previously advertised.

Similar arguments can be developed to show that it's not clear that hard incompatibilists can give compatibilist accounts of the significance of consent or the norms of paternalistic intervention, but I will not pursue that here, since they are perfectly parallel to the one sketched above. Instead, I'll conclude with the following. Hard incompatibilists like Derk Pereboom have offered powerful arguments against the idea that responsibility-entailing emotions are necessary for meaningful personal relationships. However, even if we accept that these arguments are sound, there is still reason to think that hard incompatibilism threatens the legitimacy of some core elements of our shared way of life. Hard incompatibilism may yet be true, but it is perhaps a great deal *harder* than its proponents are willing to admit.

WORKS CITED

Anglin, W.S. 1990. *Free Will and the Christian Faith*. Oxford University Press.

Ebels-Duggan, Kyla. 2008. "Against Beneficence: A Normative Account of Love," *Ethics* 119.1. pp. 142-170.

- Ekstrom, Laura. 2000. *Free Will: A Philosophical Study*. Westview Press.
- Fischer, John Martin. 1994. *The Metaphysics of Free Will*. Blackwells.
- Kane, Robert. 1996. *The Significance of Free Will*. Oxford University Press.
- Mason, Elinor. 2005. "We Make No Promises," *Philosophical Studies* 123.1-2. pp. 33-46.
- Milam, Per-Erik. 2016. "Reactive Attitudes and Personal Relationships," *Canadian Journal of Philosophy* 46.1. pp. 102-122.
- Pereboom, Derk. 1995. "Determinism al Dente," *Nous* 29.1. pp. 21-45.
- _____. 2001. *Living Without Free Will*. Cambridge University Press.
- _____. 2013. "Free Will Skepticism, Blame, and Obligation," *Blame: Its Nature and Norms*, eds. D. Justin Coates and Neal A. Tognazzini. Oxford University Press. pp. 189-206.
- _____. 2014. *Free Will, Agency, and Meaning in Life*. Oxford University Press.
- Pisciotta, Trevor. 2009. "Meaningfulness, Hard Determinism, and Objectivity," *Essays on Free Will and Moral Responsibility*, eds. Nick Trakakis and Daniel Cohen. Cambridge Scholars Press. pp. 73-89.
- Scanlon, T. M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Harvard University Press.
- Shabo, Seth. 2012. "Where Love and Resentment Meet: Strawson's Interpersonal Defense of Compatibilism," *The Philosophical Review* 121.1. pp. 95-124.
- Sher, George. 2006. *In Praise of Blame*. Oxford University Press.
- Smuts, Aaron, 2016. "Love and Free Will," unpublished manuscript.
- Sommers, Tamler. 2007. "The Objective Attitude," *Philosophical Quarterly* 57.28. pp. 321-42.
- _____. 2012. *Relative Justice: Cultural Diversity, Free Will, and Moral Responsibility*. Princeton University Press.
- Strawson, P.F. 1962/2003. "Freedom and Resentment" *Free Will 2nd Ed.* ed. Gary Watson Oxford University Press. pp. 72-93.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Harvard University Press.
- _____. 2011. "Dispassionate Opprobrium: On Blame and the Reactive Sentiments," *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, R. Jay Wallace, Rahul Kumar, and Samuel Freeman, eds. Oxford University Press. pp. 348-372.

Watson, Gary. 1987. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme," *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, F. Schoeman, ed. Cambridge University Press. pp. 256-286.

Wisdom, John. 1934. *Problems of Minds and Matter*. Cambridge University Press.

Wolf, Susan. 1981. "The Importance of Free Will," *Mind* 90.359. pp. 386-405.