

Anti-libertarianism Worth Hoping for: A Consequentialist Defense for Hard Incompatibilist Moral Responsibility Revisionism

Student Name: Willem Marijnus Smit

Student Number: 6030807

Master Program: Applied Ethics

Date: 21-6-2018

Word Count: 18413

Thesis Supervisor: Dr. Niels van Miltenburg

Second Reader: Dr. Hanno Sauer

Abstract:

If we do not have a free will, this could have significant implications for our morality. Hard incompatibilism is the belief that there is no free will, and that therefore no 'true' moral responsibility exists either. In this thesis, I will argue for the moral desirability of hard incompatibilism by arguing for its moral revisionist implications regarding our responsibility practices. I will describe a consequentialist view of moral responsibility that is not undercut by hard incompatibilism. By doing so, I will show that hard incompatibilists still have a reliable way to hold people responsible. The consequentialist view also fits well with the hard incompatibilist ideal of revisionism, as most of these good consequences are good regardless of incompatibilism, and therefore make the incompatibilist theory more morally attractive than compatibilism and even libertarianism. This will also affect our emotional reactions that are related to responsibility. Although I accept some natural basis in our responsibility related emotions, I will argue that we have good reason for a revision in the expression of our pre-reflected reactive emotions by means of a change in cultural emotional norms. Still, a reflective stance is also at all times desirable. Both of these are thus in way to approach revisionism there. I end up taking a more moderate stance than my incompatibilist predecessors, by saying that incompatibilist blame is *non*-deserved, rather than undeserved. Therefore, I do not opt for full abolishment of these reactive emotions, which I think is not realistic anyway.

Table of contents

Introduction.....	pp. 4-6.
1. Metaphysical starting point.....	pp. 7-11.
2. Moral responsibility without 'true' moral responsibility.....	pp. 11-12.
• 2.1 Two notions of moral responsibility.....	pp. 12-13.
▪ 2.1.1 Merit-based views on responsibility.....	p. 13.
▪ 2.1.2 consequentialist views on responsibility.	pp. 13-14.
▪ 2.1.3 compatibility of responsibility account...	pp. 14-18.
• 2.2 a defence for my consequentialist view.....	pp. 18.
▪ 2.2.1 practical accounts of blame and punishment.....	p. 19.
▪ 2.2.1.1 'merely a means to an end'.....	pp. 19-22.
• 2.2.1.2 Two problematic punishments.....	pp. 22-26.
• 2.2.2 Personal interactions.....	pp. 26-27.
• 2.3 Conclusion.....	pp. 28.
3. Moral responsibility and our attitudes.....	pp. 29-31.
• 3.1 Against the insulationist move.....	pp. 31-33.
• 3.2 The impossibility challenge.....	pp. 33-34.
○ 3.2.1 'Natural' emotions	pp. 34-38.
○ 3.2.2 Impossible to live in?.....	pp. 38-42.
• 3.3 The outcome challenge	pp. 42-44.
• 3.4 Conclusion.....	p. 45.
Conclusion and final remarks.....	pp. 46-48.
References.....	pp. 49-51.

Introduction

Do we have a 'free will', in the sense that *we* are ultimately able to decide what we do? I don't know. Does it *matter*, morally speaking? I think it does. Many philosophers have aimed to answer these questions over centuries. Regarding the first question, those who answer to it positively can be called 'libertarians,' and those who answer it negatively can then be called 'anti-libertarians'. Among the defenders of anti-libertarianism are those who think the universe is deterministic. I am specifically interested in causal determinism, a variant of which I will explore in this thesis. Causal determinism, roughly speaking, the belief that everything that happens is determined since it is the only possible outcome of all previous relevant events in the universe following the laws of nature. Many determinists think that it is then consistent to believe that these determined things include our thoughts and choices as well, since those are ultimately caused by natural facts about the world that we have no control over. This anti-libertarian understanding of determinism is typically referred to as 'hard determinism'. So, adhering to hard determinism leads a possible answer to the first question and, in this essay, I will assume that a variation of it is true, in order to explore the second question: "does it matter for our morality whether we have a free will or not?"

As I said before: I think it does, and so I will argue in this thesis. This means that I will be siding with the 'incompatibilists', and oppose those who say it does not matter, namely the compatibilists. I think incompatibilism is true, because, if anti-libertarianism is true, for example by means of causal determinism, it implies that people are not *ultimately* the cause for their actions, and thus also not morally responsible for their actions.¹ From this the moral claim follows that it would be wrong to hold people responsible. If 'no free will' implies no moral responsibility, this situation would appear to be very problematic for our moral practices. Are people never allowed to blame others since *all blame is undeserved*? Should we not *want* for hard determinism to be wrong, so we can have our morally desirable free will, by means of which we can keep

¹ Incompatibilism as described by Kane: "U: For every X and Y (where X and Y represent occurrences of events and/or states), if the agent is personally responsible for X, and if Y is an *arche* (or sufficient ground or cause or explanation) for X, then the agent must also be personally responsible for X." Robert Kane, *The Significance of Free Will* (Oxford: Oxford University press, 1998), 35.

these important moral practices? The compatibilist is right to point out that we have a lot to lose that we might not be able to afford to lose.

I agree with the compatibilist worry that we cannot just give up these moral practices, but I agree more strongly with optimistic incompatibilists: we should and can significantly change our practices. Therefore, some incompatibilists have argued for significant moral responsibility *revisionism*, so that our practices are more in line with a determinist, anti-libertarian worldview.² We can devise alternative practices concerning moral responsibility that do not require *desert*. If this is true, it appears that we might be able to keep most aspect of our morality. But still this would seem to be a bad thing as compared to having a free will; it would be much easier if we could just continue business as usual.

In this thesis, my first goal is to defend incompatibilist based revisionism against a variety of strong compatibilist claims that say that revisionism would either be morally problematic, or just practically impossible. However, the upshot is that I manage to give good consequentialist reasons for that we should *want* the world to be anti-libertarian, since it would open up for revisionism that would make a better world than our current libertarian one. I claim that the implications of anti-libertarianism are *morally desirable*.

To do this, I will describe a consequentialist based moral responsibility that is compatible with incompatibilism and defend this against a variety of theoretical and practical worries that could be, and have been, raised by compatibilists or pessimist incompatibilists (Chapter 2). Subsequently, I take on the most crucial challenge for incompatibilist revisionism, namely our *reactive attitudes*.³ It being thought that we, as humans, have certain emotional intuitions that are crucial for our responsibility practices. As these are typically in line with the standard libertarian or compatibilist beliefs regarding responsibility they therefore thought to be morally undesirable. Hard incompatibilist revisionists typically claim that revisionism is required in our emotional repertoire by abolishing our reactive attitudes. I will discuss this 'abolitionist' position of revisionist hard incompatibilism, and take a more modest, but also more realistic, stance than other abolitionists (Chapter 3). Before I do this, I will give a short explanation of the metaphysical position which moral impact I am going to defend. I will be modest in my

² This is most prominently defended by Pereboom in various works.

³ Peter F. Strawson, "Freedom and Resentment", *Proceedings of the British Academy*, 48 (1962), 1-25.

metaphysical claims, as this is a thesis in *ethics*, not in metaphysics. Yet, as for any discussion it is important to sketch out a clear playing field - what is being compared to what? Furthermore, I think the metaphysical position I defend needs to be at least somewhat plausible if it is to be worth considering at all, because, clearly, we need not take seriously any ridiculous ideas. Therefore, to make this defense, I will first provide a formulation of the type of incompatibilist view I want to defend (Chapter 1).

1. Metaphysical starting point

In this thesis, I will join the ranks of Derk Pereboom and Per-Erik Milam and defend the moral implications of an optimistic incompatibilist position on free will and moral responsibility, called 'hard incompatibilism.'⁴ The short answer to the question "what is hard incompatibilism", is that it is the conjunction of two concepts mentioned earlier: hard determinism, and incompatibilism. It holds that, because the world is (quasi-)determinist there is no free will, and therefore no moral responsibility exists either. In this chapter, I will try to give an elaborate answer to this question in order to have a clear image of what I understand by hard incompatibilism by quoting and analyzing Galen Strawson's Basic Argument.

First of all, hard incompatibilism requires a strong view rule of causality in the deterministic sense. I therefore assume that there cannot be such a thing as a *causa sui*.⁵ I will not claim or argue here that a *causa sui* is indeed impossible, but I rather merely state that I have thus far not found any convincing account for the possibility of there being one. I thus think this premise is at least somewhat plausible. With this assumption in hand, let us take a look at a well formulated hard incompatibilist account; the Basic Argument by Galen Strawson:⁶

1. [I]nterested in free action, we are particularly interested in actions that are performed for a reason (as opposed to 'reflex' actions or mindlessly habitual actions).
2. When one acts for a reason, what one does is a function of how one is, mentally speaking. (It is also a function of one's height, one's strength, one's place and time, and so on. But

⁴ First coined in: Derk Pereboom, *Living without Free Will* (Cambridge: Cambridge University Press, 2001), XX.

⁵ A 'causa sui' means that something 'caused itself'. To illustrate this, Friedrich Nietzsche famously described the absurdity of the *causa sui* by describing the fantastic Baron von Munchhausen who got stuck in a swamp while riding a horse through it. Dealing with this nasty situation, he pulled himself and his horse out of the swamp by pulling his own hair. This is nonsensical of course. Von Munchhausen defies gravity, he apparently creates a force that can stand apart from all natural laws that would cause this action impossible. Friedrich Nietzsche, *Beyond Good and Evil*, (1886), 21.

⁶ Galen Strawson, "The impossibility of moral Responsibility", *Philosophical Studies* 75.1-2 (1994): 13-15. Based on this argument, Strawson argues that moral responsibility is impossible. I assume that it at least makes it *implausible*, given I am right to assume the implausibility of there being a *causa sui*.

the mental factors are crucial when moral responsibility is in question.)

3. So if one is to be truly responsible for how one acts, one must be truly responsible for how one is, mentally speaking - at least in certain respects.
4. But to be truly responsible for how one is, mentally speaking, in certain respects, one must have brought it about that one is the way one is, mentally speaking, in certain respects. And it is not merely that one must have caused oneself to be the way one is, mentally speaking. One must have consciously and explicitly chosen to be the way one is, mentally speaking, in certain respects, and one must have succeeded in bringing it about that one is that way.
5. But one cannot really be said to choose, in a conscious, reasoned, fashion, to be the way one is mentally speaking, in any respect at all, unless one already exists, mentally speaking, already equipped with some principles of choice, 'PI' - preferences, values, pro-attitudes, ideals - in the light of which one chooses how to be.
6. But then to be truly responsible, on account of having chosen to be the way one is, mentally speaking, in certain respects, one must be truly responsible for one's having the principles of choice P1 in the light of which one chose how to be.
7. But for this to be so one must have chosen P1, in a reasoned, conscious, intentional fashion.
8. But for this, i.e. (7), to be so one must already have had some principles of choice P2, in the light of which one chose P1.
9. And so on. Here we are setting out on a regress that we cannot stop. True self-determination is impossible because it requires the actual completion of an infinite series of choices of principles of choice.

10. So true moral responsibility is impossible, because it requires true self-determination, as noted in (3).

In this thesis, I will assume that a hard incompatibilist account based on something like the Basic Argument is true. I will, therefore, break down shortly below how I think the argument ought best be understood and describe my hard incompatibilist position in more detail.

First of all, step 1 and 2 discuss 'free' action, which means *acting on reasons* specifically. A reason being defined as "a function of how one is, mentally speaking". Therefore, by accepting the Basic argument, I also take it we are still *rational agents*,⁷ this includes being *moral agents*.

Moving on in the argument, step 3 and 4 shortly explain the *incompatibilist* view: for an agent to be truly morally responsible, she would need to be the true *source* of her actions in order to be *deserve* blame for this. 'True self-determination', can be thought of as Strawson's formulation of the free will. It would allow one to be the source of their actions and thus be truly responsible for them. But if a deterministic account (like the one that is given in steps 5-9) is true, then there is no true self-determination and therefore step 10 follows: there is no *true* moral responsibility.

Next, step 5 accompanied with steps 6-9, is the hard determinist explanation leading to hard incompatibilism. If this causal determinist explanation is correct, we thus cannot be the ultimate cause of our mental states, because these mental states are caused by other of our mental states and so on; until we can only say that there were events clearly outside of our control, either social/psychological or biological, that have caused these deepest mental states which ultimately cause all the others. If this description is correct, and considering my premises, there cannot be a free will, and therefore no true moral responsibility.

Now to give a scientifically accurate account I need to make one small swerve away from causal determinism: quantum mechanics states that, at the smallest knowable level, it appears to be the case that there is no causal determinacy.⁸ As far as our knowledge currently goes, electrons merely have a *chance* to be at a certain place at

⁷ As is also affirmed by Pereboom to be crucial for the formulation alternative responsibility strategies. Derk Pereboom, *Free Will, Agency, and Meaning in Life* (Oxford: Oxford University Press, 2014), 136.

⁸ Pereboom "Living," XVIII.

a certain time; events are not determined, but essentially *random*, at the smallest level. Although this suggests determinism is essentially false, and thus things are *indeterminist*, this does not imply in any way that people could have free will. The indeterminism I will defend still holds on to strong causality as much as causal determinism. I thus hold that this version of indeterminism does not allow for a free will either, so this does not affect my theory. We might say things are to us 'quasi-determinist'.

In this chapter, I have given an explanation of my hard incompatibilist position by describing the Basic Argument, most importantly that we can still give, and act on, reasons. I think this is a plausible hard incompatibilist account against libertarianism, and thus against true moral responsibility as well. I do not dare to claim that this metaphysical account is the *most* plausible one, as much more in-depth accounts have been given for and against this view, but that was also not my goal here. My goal here was merely to formulate a plausible account of which's moral implications it makes sense for me to discuss.

In the next chapter, I will try to explore the question where the hard incompatibilist account as just described here would leave us regarding moral responsibility. To which the answer will be: at a much better place than one would think.

2. Moral responsibility without 'true' moral responsibility.

Where does a hard incompatibilist account, such as the Basic Argument, leave us concerning moral responsibility? The incompatibilist view typically holds that it is *always unfair* to hold one responsible, because people never truly *are* responsible. Therefore, they cannot truly deserve blame or praise. In this thesis, I will be specifically concerned with blame.⁹ To prevent this unfair blame, we need *revision* in our institute of moral responsibility. Compatibilists on the other hand, tend to think that even if the Basic Argument is valid, our current moral responsibility practices are still justified. In this chapter, I will make a defense for how the hard incompatibilist can plausibly deal with removing true moral responsibility from our moral practices and still have a well-functioning moral system. By doing so, I will argue that incompatibilism does a better job at this than compatibilism, and secondly, that the resulting situation would be *preferable* over our current situation, at least for the consequentialist.

Regarding the implications for moral responsibility, the Basic Argument shows that there could be no true moral responsibility without one being the ultimate source of one's actions, but it does not necessarily hold that there could be no justified form of moral responsibility at all. The intuitive, and commonly held, notion of responsibility, namely that someone *is* responsible, in virtue of one's free will, is indeed denied. Yet, I think that in an anti-libertarian society we could still *hold* someone responsible, based on a different fundament. This begs the question: "would this then not be compatibilism?" Not necessarily, as I will show that the consequentialist form of responsibility I defend still opts for a significant amount of *revisionism* regarding our moral responsibility practices, specifically in our justice system, and that it is therefore still rather incompatibilist than compatibilist.

Therefore, my approach here was to essentially ask a Dennettian kind of question: "what kind of [moral responsibility] is worth wanting?"¹⁰ However, I will stay more true to general semantics than Dennett and will not try to attach "moral

⁹ For an optimistic account of hard determinism concerning *praise*, I advise you to look up Saul Smilansky. Smilansky argues for acting morally, *for only moral reasons*, not for any expected praise. One acts more 'purely' for what is good than for one's pride. Saul Smilansky, "The Ethical Advantages of Hard Determinism", *Philosophy and Phenomenological Research* 54, no. 2 (1994).

¹⁰ Daniel C. Dennet, *Elbow room: The varieties of free will worth wanting* (Boston: MIT Press, 1984).

responsibility” to something that is essentially something else, like he did with “free will”. I will be especially keen on separating the supposedly undercut true moral responsibility from other kinds of moral responsibility. Still, I think I can make a good account of what moral responsibility is worth wanting: A non-standard account of moral responsibility, namely one ultimately based on *good consequences*. I will defend this account against criticisms of varying nature. By doing so, I mean to create an even clearer image of what kind of consequence based morality can be accepted by the hard incompatibilist, and what kinds of revision it actually requires.

2.1 Two notions of moral responsibility

As moral responsibility is probably the most important way in which free will affects our morality, it is important to understand its role in our morality. Moral responsibility is a fundamental part of our moral practices. It allows us to meaningfully say that people have certain moral obligations, certain ‘oughts’ on which we can expect them to act, such as following a variety of moral norms. Disregarding certain exceptions, we generally think of people as rational and moral agents, and therefore constrained in their actions by these obligations as it gives them reasons to act in a certain way. Based on how one acts (or intends to act) in light of these moral obligations, we think it is justified to blame or praise that person.

For example, when we say that a tornado is ‘responsible’ for something, say, a lot of property damage, we tend to think this is fundamentally different from a situation where *I* am responsible for a lot of property damage. While both of us could be rightly thought of as being the entity that was physically dealing the damage in what has been a regrettable event, but we believe that only *I* had a moral obligation, a responsibility, not to do so. We believe that, since *I* am considered a rational moral agent, *I* am therefore in the position where we could say that *I* have moral reasons to not inflict any damage on anyone else’s property. A tornado, on the other hand, is not a rational agent and will thus not act on reasons. Why my rationality allows me to be rightfully blamed in this situation, as opposed to the tornado, is justifiable through different methods of reasoning.

The main approaches of giving justifications for moral responsibility date back to Aristotle's *Nicomachean Ethics*.¹¹ He discusses two different fundaments based on which one might attribute responsibility to someone, a *merit-based view* and a *consequentialist view*. In this section, I will show that the commonly practiced merit-based view is undercut by the Basic Argument, but that the hard incompatibilist still could, and should, take on a consequentialist view on moral responsibility.

2.1.1 Merit-based views on responsibility

When people talk about moral responsibility, they usually think of a responsibility that is merit-based. Merit-based views place the basis of moral responsibility on some quality of the agent. In virtue of possessing this quality or 'merit', she would be *deserving* of certain blame and praise as a response to her actions. I think this is quite intuitive belief, and therefore likely emotionally backed.¹² As discussed, the most prominent merit on which responsibility is based is 'free will' since it would make one, in a way, the ultimate source of one's actions, as one would have the ability to do otherwise, and therefore *truly* responsible. I would be deserving blame when I intentionally caused property damage if indeed my choice to cause this damage is truly my own. I am rightfully blamed, because the actions is my own. I could have chosen to do otherwise, but did not. It is thus exactly this kind of moral responsibility that is described in the Basic Argument as being *true* moral responsibility, and whose existence is being challenged by it, and therefore, of course, the kind whose moral relevance I am significantly challenging in this thesis.

2.1.2 Consequentialist views on responsibility.

Alternatively, consequentialist views on responsibility suggest that the blame and praise need to be attributed in such a way that would lead to desirable consequences, instead of it being truly deserved. Of course, there are certain features (rather than merits) required for an agent to be susceptible to praise or blame. To use the tornado example again: Blame would not affect the tornado's harmfulness at all (one would aptly

¹¹ Aristotle, *The Nicomachean Ethics*. Book III, 1-5.

¹² Ralph Wedgwood, "Moral Disagreement among Philosophers", in, *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Michael Bergmann & Patrick Kain (eds.) (Oxford: Oxford University Press, 2014).

be called crazy for trying to do so), while I could take the praise and blame from my peers as *valuable moral judgments*. I take it as a form of feedback based on which I could improve my behavior, as it was not in line with certain norms. Even more so: if I would not receive any praise or blame regarding my actions, I could have trouble knowing which actions are morally desirable and which are not. But keep in mind that this praise and blame would, of course, not be *deserved*, but useful for good consequences. Next to these moral reasons, we can consider *prudential* influences as well when concerning consequentialist views on responsibility. This is generally in the form of reward and punishment, like a good salary on the one hand, and a fine or jail sentence on the other. Additionally, one could be praised for their good behavior and feel *proud* because of this.¹³

Consequence based moral responsibility can base blame and praise on a variety of factors, in order to assess whether one can rightly be *held* responsible. This could mean that either we can question an agent's motivations, and, when necessary for expected better future behavior, blame and punish them. We could also help them become better moral agents in any other way. In 2.2.2, I will discuss some of these consequentialist methods in more detail, as to show that is this form of moral responsibility need not to be confused with 'corrigibility'.¹⁴

2.1.3 Compatibility of responsibility accounts

Although 'good consequences' is a highly abstract term that is open for interpretation, opting for good consequences is a moral norm that can be recognized by almost anyone. So, interestingly enough, some compatibilists, like Shaun Nichols, argue that our current intuitive responsibility practices are justified *for their good moral consequences*.¹⁵ Our current intuitive merit-based view on moral responsibility seems to

¹³ But of course, our pride might be undercut by hard incompatibilism. We cannot truly take pride in our achievements as much as we cannot be for their wrongdoing. Still, I think people can be happy with their lives as they are.

¹⁴ Taylor discusses an uncharitable view of determinist moral responsibility, in which man is not to be seen as 'rational', thus leaving open only rewards and punishments, as if we are mice that are to be taught by means of conditioning. I thus think this is not in line with my view, as I do take man to be rational. Charles Taylor, "Determinism and the Theory of Agency", in *Determinism and Freedom*, ed. Sydney Hook (1959), 211-18.

¹⁵ Shaun Nichols, "After Incompatibilism: A Naturalistic Defence of the Reactive Attitudes", *Philosophical Perspectives* 21 no. 1 (2007), 405-428.

function quite well in achieving good cooperation.¹⁶ If this is true, then why would we ever want revisionism? We still want revisionism for two reasons, one being that this belief that they tend overlap is *mistaken*, for empirical reasons, part of which I will discuss here, and another part in the next chapter. The second is based on a hard incompatibilist belief, namely that it is at all times *unfair* to blame someone.

Considering the first, there is specifically one thing based on which these two responsibility views tend to conflict, which is on what they think is the correct moral justification for blame. This crucial difference is that consequence-based views tend to be at all times *forward-looking*. Simply put, this means that if we punish someone, this has to be done in such a way that this will lead to better overall outcomes in the future (which is, of course, generally better moral behavior as performed by the punished agent). Oppositely, merit-based views rather have us punish based on the wrongness of the act. Meaning that they react on how good/bad one's will was when performing this action and a fair praise/blame and reward/punishment is in proportion to the goodness/badness of this will. It can be said that it is *backward-looking*, as it 'looks back' to that specific action and reacts to just that action performed by that agent in a way that is appropriate. Other factors do not matter for blame based on merits. Unlike forward-looking responsibility practices, this need not necessarily reap the best consequences, as its justification lies elsewhere.

The forward-looking aspect of consequence based responsibility practices means that if we give up the idea that responsibility is deserved, ideal punishments would almost become an empirical question.¹⁷ It then happens to be the case that, at least some, empirical data suggests that our intuitive punishments tend to be much harsher than what would result in optimal consequences. Shorter jail sentences often reap better effects on improving one's moral character than longer ones, and it appears that many jail sentences highly overshoot this optimal sentence and even tend to *worsen* many criminals moral character, making them more likely to inflict harm in the future.¹⁸ Instead, it is suggested that positive influences rather than punishments often do a much

¹⁶ Nichols, "After Incompatibilism", 417-419.

¹⁷ Neil Levy, "Less Blame, Less Crime? The Practical Implications of Moral Responsibility Skepticism." *Journal of Practical Ethics* 3 no. 2 (2015): 14.

¹⁸ Neil Levi. "Punishing the Addict: Reflections on Gene Heyman," in *The Future of Punishment* (2013): 233-245.

better job improving one's moral character.¹⁹ This suggests that our merit-based intuitions of blame are suboptimal for reaching good consequences.

This outcome seems quite plausible to me, as it would be *extremely coincidental* if both responsibility views would happen to end up opting for exactly the same kind of punishments at all times. So, even if these empirical claims on responsibility practices were, in fact, mistaken, this would not deny the likelihood of a clash between the two practices. If typical merit-based justifications for these backwards-looking responsibility practices would be invalid, as the hard incompatibilist position proposes, we would be free to refrain from considering the backwards-looking punishment, and only need to look forward, and thus exclusively aim at achieving good consequences.

The second reason why revisionism is implied by following the consequentialist view is because it is commonly thought by incompatibilists that it is *unfair* to blame one. One could find it quite striking then that the consequentialist responsibility justification is presumed to be compatible with an incompatibilist view, as justifiable moral responsibility is thought to be *the* topic of discussion of the compatibilist-incompatibilist debate, where the compatibilist usually affirm and the incompatibilists deny its possibility. Here thus is exactly where I start to disagree with my predecessors. For incompatibilists, following a consequentialist account of blame would still come down to blaming people who we are essentially not morally allowed to blame, as it would be unfair. Therefore, it seems that the notion of responsibility I argued for might actually be compatibilist, which means that I will be acting unfairly when I keep blaming people, but as this is important to do for good consequences, it might be good to do so regardless. This is almost exactly what Nichols concludes after weighing these matters off in a reflective equilibrium.²⁰ I find this conclusion unsatisfying because (a) he seems to just take the unfairness for granted, and (b) too arbitrarily assumes that revisionism is undesirable by relying on questionable research regarding moral emotions.²¹ I thus disagree with him on each of these points.

First of all, I am much less skeptical, even rather optimistic about revision in our responsibility related emotions. But, as I have mentioned, I will discuss this extensively in chapter 4. Second, I disagree with the standard hard incompatibilist view on blame

¹⁹ Mark Kleiman, *When Brute Force Fails: How to have Less Crime and Less Punishment*. 2009.

²⁰ Nichols, "After Incompatibilism", 423-24.

²¹ I will discuss this in detail in 3.3.

and fairness in an incompatibilist world. For hard incompatibilists it appears to be the case that, as the merit that allows one to be rightfully blamed is missing, it is at all times unfair to blame someone, since people are never deserving of blame. But I think that if all blame would be unfair then the word 'unfair' becomes *redundant*. I think for us to rightfully say that some blame is 'unfair blame' there needs to be some 'fair blame' for it to be in contrast with. However, there is no such fair blame. I thus argue that, in this case, talking about 'unfair' as applying blame that is not 'deserved' is just as absurd as talking about 'theft' in a world without the concept of 'property'. I think that when blame and fairness are understood this way, blame is simply criticism and possibly a *harm* (if it comes with a punishment) that needs justification.²² If we take out the fundament of all deserved blame, this is what the hard incompatibilist should end up with. So, when we blame someone there is a harm, and it is generally agreed that harming someone is *prima facie* wrong. Therefore, before we blame anyone, especially through punishment, this needs good epistemic justification. As merit-based blame is intuition-based and thus likely to get suboptimal moral consequences, we think that the required justification cannot merely be an intuition. This means that if we were to continue 'business as usual' people are going to be harmed in ways that lack proper justification. For optimal consequences, revisionism ought to have us change our blaming practice to one in which we are much *more sceptical* of- and thoroughly reflect on our intuitions, which might have us take an 'objective attitude'²³, which I will discuss in detail in the next chapter. As such, I think Nichols is mistaken.

If my description is correct, it does indeed appear we could have a functioning moral responsibility system based on other grounds than a free will. Given anti-libertarianism, we need good alternative justifications for our blame, which are not given by holding on to business as usual, since this is essentially the merit-based practice that has lost its most important fundament. Instead, we rather need to reflect well on our intuitions, for their good consequences, instead of taking them for granted. Therefore, it appears that the incompatibilist revisionist can give better justifications for the consequentialist blame than the compatibilist. Finally, there is also an argument

²² One can even go so far as to say that all blaming attitudes are already harming that person, as you put him in an position of decreased respect. Even only mentally, he is already being condemned and put in the category of wrongdoers. An unfair form of ceasing basic respect, similar to slander. Per-Erik Milam, "Reactive attitudes and Personal Relationships", *Canadian Journal of Philosophy*, 46, no. 1. (2016): 115.

²³ Strawson, "Freedom and Resentment", 6.

against the libertarian. Based on the forward-looking / backward-looking distinction of blame justification, it is the case that, for the hard incompatibilist account and not for the libertarian account, the justified blame is the very same as the blame that has good consequences. Libertarians have a conflict of justified blame, that anti-libertarians (given they are *revisionist*) do not have.

2.2 A defense for my consequentialist view

I now hope to have now convincingly shown that hard incompatibilism does do have access to a method of justifying responsibility, unlike what is often believed to be the case. I do think that this responsibility account still begs a few questions. In this thesis, I will try to elucidate my view on consequentialist responsibility that is compatible with hard incompatibilism by defending it against some skeptical worries. To make a sufficiently strong case for the moral desirability of consequentialist responsibility, I think that I will have to deal with specifically the two following challenges:

- It appears to enable punishments that might be very much against our normative beliefs. In this subsection, I will make a defense against some practical worries on moral responsibility (2.2.1);
- Our emotional reactions would appear to be unreliable, which suggests *emotional revisionism* as well, which is potentially problematic. (2.2.2; more thoroughly in chapter 3).

I think it is important for the moral desirability of the position that I am defending that all of these issues are to be dealt with. Unfortunately, this is not the time and place to completely root out all these worries, but I will make effort to give responses that are sufficiently satisfying for the time being, so that these worries will not turn into problems that manage to undermine my theory. If I am successful, I think my hard incompatibilist case will be sufficiently strong as to create a convincing revisionist imperative.

2.2.1 Practical accounts of blame and punishment

The next challenge for my defense for on hard incompatibilist moral responsibility is a series of practical moral problems. These worries were raised by Scott Sehon against specifically Pereboom's account.²⁴ As it is the case for the consequentialist responsibility view that blame and praise always require justifications other than that the agent 'simply deserved it', there are especially two practical problems Sehon has regarding the implication of the implications of Pereboom's hard incompatibilism for our justice system. I take it that these issues affect my account as well, and thus take on the challenge to deal with them.

First, there might be a significant the hard incompatibilist way of justifying blame, as it allows us to go against the Kantian categorical imperative of treating people as 'merely a means to an end'.²⁵ Second, it appears that the hard incompatibilist's methods for punishments can lead to punishments that are by, the looks of it, extremely unjust. In this subsection, I will show that these challenges are not be as problematic as they appear, because, using consequentialist 'blame', options are quite versatile.

2.2.1.1 'Merely a means to an end'

Starting with the former problem, the worry of 'using one as merely a means to an end' is the direct negation of the second formulation of the categorical imperative, as formulated by Immanuel Kant:

"Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end".²⁶

I do not accept the *never* in this imperative for two reasons. First, because I think that to this day there is no closing proof that Kantian constructivism is correct. Therefore, I think this principle lacks a metaethical basis that would allow us to make such a claim.²⁷ Second, I think it is morally correct to torture one person, and thereby using him as

²⁴ Scott Sehon, *Free Will and Action Explanation*. (Oxford: Oxford University Press, 2016), 7-22.

²⁵ This worry was already expressed by Pereboom in "Free Will, Agency", 164.

²⁶ Immanuel Kant (1993) [1785]. *Grounding for the Metaphysics of Morals*. P. 36. 4:429.

²⁷ Famous attempts include: Christine Korsgaard's *Sources of Normativity* (1996), and Alan Gewirth's *Reason and Morality* (1978). Yet, neither is thought to be successful in this.

merely a means, to prevent the even more painful torture of a thousand others. I think there are reasonable limits to its demands. Of course, this is highly exceptional and unrealistic, and I will take it that it is still a *very strong moral principle* for respecting the value human dignity. A theory that would allow one to go against this would have a lot of moral explaining to do. However, I will argue that the account I defend usually doesn't violate it and secondly, when it does so, there are valuable norms on the line, so it will be justified.

It is to be feared that any kind of punishment that is harsher than is effective for the agent's own moral character formation would essentially be partly grounded as using this person as a means to achieve social order, by using his punishment as a way to deter others.²⁸ Against this worry, I have two rebuttals. First, it could be argued that given that we punish people when needed for the sake of a desirable society, this positively affects them, and everyone they care about as well. For each agent's ends, it is important that we have a flourishing society. Thus we hardly ever hold someone responsible as 'merely a means to an end', because the end is also part of that person's end as well, given he is not a psychopath. The person punished should, even if he doesn't, appreciate the fact that society's norms are good and that they are well protected (given this is the case, otherwise I find it much less acceptable).

Secondly, there is ample literature concerning certain responsibility concepts that put heavy restrictions on allowing people to use others as merely a means to an end. As is explained by Pereboom, these concepts are thought to be compatible with an incompatibilist view on moral responsibility, as they do not require true moral responsibility.²⁹ Among these are *accountability* and *attributability*, which can be combined in a supervenient concept: *answerability*. Of course, there are also other accounts as well (e.g. conversational responsibility, take-charge responsibility), but I do not have the space and time to discuss all of these and therefore focus on answerability specifically.

Attributability holds roughly that, if we judge person A's actions, who are a result of person A's value judgments, then it makes sense to say to address the reactions on these actions to person A and not person B. Addressing blame for A's action on B, who is

²⁸ Sehon, "Action Explanation", 8-9

²⁹ Pereboom, "Free Will, Agency", 136-41.

not in any way linked to that action, is in a practical sense, nonsensical for rational agents. Then, when regarding an agent as being *accountable*, we say she has certain obligations, for which she has to be able to give an *account*. Meaning that for reasonable discourse, she ought to be able to give reasons for his actions based on which we can identify the justifiability of our blame. Answerability then holds that we can ask agent A, that has performed certain actions, to *justify* her actions. We ought not simply blame people right away if the actions they performed appeared wrong from only our point of view, she has to have the option to defend herself. Whether the agent requires a certain form of feedback depends on her reasons for action. One's moral improvement can be reached by uncovering one's reasons by means of interaction.

It is thought that this account of answerability in responsibility makes sense to properly addressing blame to the person who has done the harm, instead of blame being arbitrarily appropriated. It is crucial here that the agents whom we might blame have a 'reason responsiveness', as they can effectively take this blame as informative for their future behavior, thereby perhaps even improving their moral character.³⁰ But note that still this only grants them the ability to be a sensible target of blame as response to an action which they have performed, not the ability to be *deserving* of blame.

Furthermore, there might be a question why it would make for bad consequences to disregard these principles. Why this principle of answerability is especially important for good consequences is because it allows us to be able to expect what is to happen to us; we can count on certain things, as there are clear norms that one has followed or not. If we disregard such principles this could lead to a society in which one might not *feel safe* in. One could arbitrarily be used for any kind of end. Not following answerability raises the well-known utilitarian hazard of the transplant surgeon case:³¹ simply maximizing consequences in short term could lead to very undesirable social effects over a longer period of time. We would be much more reluctant to visit a hospital if there is always a chance of being cut up for organ harvest. The consequence is that we would not feel safe, as it appears that we do not have any norms or rights that we can count on. I think this same argument works on the worry regarding 'punishing' innocent

³⁰ Pereboom, "Free Will, Agency", 136

³¹ In this scenario a homeless man, or a delivery man is brought into the hospital and 'cut up' by a doctor. Five of his organs are taken and used to save five other people's lives. Judith J. Thomson, "Killing, letting die, and the trolley problem." *The Monist* 59.2 (1976): 204-217.

people as a *deterrence*. Answerability therefore gives a strong rule consequentialist constraint. A norm, perhaps even a right, that we can rely on is very valuable in society against arbitrary harm. There are therefore good consequentialist reasons why we need a significant epistemic requirement for going against the categorical imperative.

2.2.1.2 Two problematic punishments

It could be worried that in more practical cases still, like criminal justice, it might be harder to avoid using certain people as merely a means to an end, for the sake of society. As criminals are never really deserving of their punishments, we have to give really good justifications for why we are punishing them. I think it greatly matters exactly *how* we do this.

A well-known account on criminal justice in hard incompatibilism comes from Derk Pereboom's *Living Without Free Will* (2001). He describes ways of 'punishing' that does not involve any kind of blame at all. For example, people could be given moral education, to make them more capable moral agents that are more likely to make good moral decisions in life.³² Alternatively and more short term, it could be possible and of good consequences to 'quarantine' criminals, analogously with people with dangerous transmittable diseases like Ebola, who are also not responsible for the harms they inflict.³³ The result is still that they are locked away from society, but phrasing it as quarantine is clearly much less blame laden than a "jail sentence". Against this account by Pereboom, and specifically his proposed use of quarantine, Scott Sehon has expressed some worries that I will now try to refute. Thereby I will show that the hard incompatibilist responsibility account is versatile in dealing with blame.

The first hazard here is that in consequentialist view, people might receive harsher punishment for committing crimes that are less severe than others that have committed more severe crimes. Some crimes are statistically generally a one-time offence, and are thus hardly suggesting any future crimes. Murder is thought to be such a crime.³⁴ Car theft, on the other hand, generally precludes more crimes. It appears then that the car thief is a higher (current) risk for society than the murderer.³⁵ Forward-

³² Pereboom, "Living Without", 161-66.

³³ Ibid, 174-77.

³⁴ Sehon, "Free Action", 12.

³⁵ Ibid, 10.

looking punishment would then likely suggest more punishment for the thief than for the murderer, for whom punishment might even be superfluous.

In a second hazard, this can be stretched even further. It might be the case that people who have not committed a crime (yet) could just as well be quarantined, since instead of actually having performed a crime justifies one of blame, it is especially important that we have reason to think they are currently dangerous. Pereboom argues that punishing these individuals would require a 'higher epistemic bar'³⁶, as opposed to criminals, of whom we already know for sure are capable of committing such crimes as they already have. But according to Sehon this higher epistemic bar might be a *farce*, partly because of what I have just described on murderers.³⁷ Furthermore, he describes that with certain relevant social and psychological studies of the population can have us assess risk groups in which there is a significant calculated risk that they will commit crimes in the future, like the 'heat list' that the Chicago police has.³⁸ This risk is expected to be much higher than that of a one-time murderer. It appears that we would then have consequentialist reasons to quarantine certain innocent people than certain murderers, which would be odd to say the least. Even though these worries are quite legitimate on Pereboom's account, I think hard incompatibilists can defend this simply by giving alternative consequentialist solutions. I will first deal with the latter hazard.

One way to determine these alternative approaches, is by trying to target important aspects of their moral character formation, or in other words to ask and deal with the question: "How is it that they are who they are?" We know that the quasi-determinist answer to this question is: ultimately due to their genes and their environment. They must have been very unlucky to have become bad moral agents, in whatever way it happened. Then if we punish them harshly on their immoral character, they end up being *double unlucky* in their lives.³⁹ First, they have had, as is thought to generally be the case, a tough and stressful childhood, and then they would be locked up for having become dangerous citizens because of this. Especially since it is often partly a social problem, it is important to look at other solutions than adding more harms, if possible. We might be able to deal with the factors that cause their moral characters to

³⁶ Sehon, "Free Action", 12.

³⁷ Ibid, 13.

³⁸ Ibid, "Free Action". 12-13

³⁹ Levy, "Less blame", 11.

develop poorly and prevent one from becoming criminal quite early on. I think that this is much rather the apt way to deal with the criminals: the path to success is rather to be expected in on the *social factors*, like poverty, that make one a criminal.⁴⁰ This means that the consequentialist view would suggest, something very much different for the Chicago 'heat list' than Sehon fears. If anything, the potential criminals from Chicago needs improved social conditions and moral education as to enable both their and society's flourishing, if it is not too late. If it *is* too late, (which I do think is quite hard to assess) then I think quarantine *can* be the answer. Sadly, this means that the person's moral character has been formed to such a state where interacting with the agent can almost not be anything but harmful. I would thus suggest that it is best that the agent ought to be kept away from society in a respectful way, but perhaps not too comfortably still, for we do not want to motivate people to get diagnosed as moral insane by committing crimes on purpose.⁴¹

Secondly, regarding the issue of how to deal with justice regarding murderers who are statistically one time offenders, I think it is in this case, again, not the best option for the consequentialist to 'punish' by means of quarantine as it is clear that it is not too effective. Again, it is important to understand that the consequentialist view on responsibility can offer versatile solutions, and I will therefore show how this can be done in this practical challenge just like in the previous one. When considering how to deal with murder and murderers, who are not necessarily in need for character improvement or quarantine, I think that the best consequences are gained by making an effort to ensure proper expression of important norms and values that are good for a society to hold. Instead of merely to 'keep wrongdoers out of society for a certain period', expected punishments are also a good method to *deter* people from committing crimes. These punishments ought not so much be seen as a reaction to the specific wrongdoers for the character formation, but should also deter potential other wrongdoers by communicating a moral norm of the society (giving a moral reason) and a *prudential reason* for people to not cross it. Therefore, because murder is a type of crime that is quite universally accepted as horrendous and extremely undesirable in to

⁴⁰ Christopher Lyons & Becky Pettit. "Compounded Disadvantage: Race, Incarceration, and Wage Growth." *Social Problems* 58 (2011): 257-280.

⁴¹ Saul Smilansky, "Hard Determinism and Punishment: A Practical Reductio." *Law and Philosophy* 30 (2011): 353-367.

have in any societies, it would require a quite heavy punishment as a promised deterrence since it is so undesirable. It wouldn't be, it could be expected that less empathetic people might kill more often. I think this is what Sehon aims to point out: there is a significant discrepancy in how criminals should be treated as to deter people from violating important these norms and values, like the undesirability of murder, and Pereboom's approach of quarantining and educating people as to not hold them responsible in any way. Now, are these people then partly being used as a means to an end? I think partly yes. But I think it can be sufficiently justified, though it might not be all too clear how exactly to balance sufficient deterrence that our societal norms are protected and we use people as a means as minimally as possible, and as respectfully as possible.

It becomes clear by dealing with these worries that Sehon raises that there is a difference between the two ways of punishment. Deterrence works best with very heavy punishments, just imagine the drop in the amount of people speeding if one would receive the death penalty for it, while opposingly, it is thought that positively affecting one's moral character requires punishments are that are not that harsh. In cases like murder it is clear that deterrence ought to do most of the of getting good consequences, and in the case of young potential future criminals, moral education is more useful. A proper balance has to be found in each scenario.⁴²

If we then relate this back to the worry of 'using one as merely as a means to an end', the approaches where deterrence is needed does not fit with the principle of answerability, since part of the blame is not being attributed for character formation. Deterrence *can* be based on good consequences, but not on answerability. Thereby there is a part of 'using one as merely a means to an end' then after all, but I think that this is not too problematic, because I think I have given a sufficiently good moral explanation for this, namely that murder is something we very much would like to prevent in our society. For its good consequences, we are allowed to put up some deterrence in a hard incompatibilist world, but always need to weigh this out against the strong Kantian principle. I think this is a reasonable challenge raised by Pereboom and Sehon, as it might make the hard incompatibilist position unappealing to the Kantian.

⁴² Levy, "Less Blame", 12.

2.2.2.2 Personal interactions

As I have mentioned before, the merit-based view more strongly corresponds with our intuitions than our consequentialist view of responsibility. It appears to us that someone *is* responsible for their actions, not that we merely *hold them* as responsible. However, when looking at the biological grounding of this intuition, it is found that many other social animals tend to hold each other responsible too.⁴³ This strongly suggests that the practice cannot be not originally founded our 'distinctively human' capacity of reason, but rather some given intuitions as a social animal. Which in turn raises strong doubts about their expected validity, since the Darwinian Dilemma suggests that we have no ground to believe that these natural intuitions are hinting towards any moral truth.⁴⁴ Instead, the only thing we might be able to say about some of these intuitions is that they are universal across almost all cultures,⁴⁵ which means that this tendency is *natural*. And it is thought that, as an emotional reaction of holding one responsible, all cultures have anger. But exactly *how* this anger is expressed differs significantly.⁴⁶ Western expression of anger can be described as a retributive desire that causes people to respond by blaming and punishing those that have caused us to experience this anger.⁴⁷

Like I already mentioned in 2.1.3, we tend to achieve suboptimal consequences when unreflectively acting on these emotion based intuitions. Therefore, we ought rather take a reflective stance to them at all times, like an 'objective attitude'. This can be contrasted with the reactive attitude. Reactive attitudes are the emotions we have that define our intuitive moral responsibility. Reactive attitudes are thought to include the following emotional reactions: anger, moral indignation, resentment, guilt and gratitude. It is important to stress that as soon as we take on a reactive attitude we view this person as responsible. In a sense, we immediately start to blame (or praise) them, mentally. So, to completely root out false blame, it might be the case that we have to

⁴³ Jessica Flack & Frans De Waal, "Any Animal Whatever", *Journal of Consciousness Studies*, 7, no. 1-2 (2000).

⁴⁴ Street, "Darwinian Dilemma", 115.

⁴⁵ Street, "Darwinian Dilemma", 115.

⁴⁶ Owen Flanagan, *The Geography of Morals: Varieties of Moral Possibility*. (Oxford: Oxford University Press, 2016).

⁴⁷ Shaun Nichols, "After Incompatibilism", 412-413

completely abolish these reactive attitudes. This hard incompatibilist revisionist position is called 'abolitionism'.⁴⁸

As I also mentioned in 2.1.3, most some hard incompatibilists think that taking on such a reactive attitude emotion would already be morally wrong in itself, because it already *is* a blame, which is *unfair*. Instead, I think it is wrong because taking such an attitude is to unreflectively blame, and thus to often *harm while lacking a good justification*. Though this is not much of a difference, I think my account is a less resolute on the taking such an emotional reaction and therefore less demanding of to have such a reaction. This has a positive and a negative impact on for my case, as compared to my predecessors. The positive impact is that if we do not manage to abolish all of these reactive attitudes this would be less of a problem than it would be according to other abolitionists. The negative consequence is the other side of the same coin: we have less reason to make the change and this lessened importance of the change makes it a less strong reason for revision, and thus makes it also less likely to happen. Still, I think it is more realistic and thus *more* like to happen, as total extirpation of the reactive emotion is probably impossible.

This leads us to the view that it would not only be the case that we have reasons to change very practical aspects of our society, like our justice system, but even some of our emotional reactions as well. This thus seems to be going much deeper into our core as human beings. But I will argue that it is, in fact, much more a cultural thing. Still, some have argued that revisionism in our emotional attitudes is (nearly) impossible, as well undesirable since they play crucial roles in different parts of our lives; it even is suggested that it is these attitudes that make our moral responsibility, more so than being the source of one's actions. The issue of the attitudes has famously been raised by Peter Strawson in his 1962 landmark article: *Freedom and Resentment*. As the hard incompatibilists are the ones opting for a change, the burden of proof is on us. In the following chapter, I will discuss these attitude related challenges raised by Strawson in detail, as it is paramount that they are properly dealt with for the incompatibilist standpoint, and I will show we can do so very well.

⁴⁸ Note that there are some *positive* reactive attitudes as well. It is often held against the abolitionist that taking on the objective attitude would have us abolish genuine gratitude as well, because also no one is truly deserving of this gratitude (E.g. Sehon, "Free Will", 18) However, I think Tamler Sommers has already given an apt defense for it, and I would thus advice one to simply find this. I thus take it he has properly dealt with this issue and will not discuss this issue any further here.

2.3 Conclusion

In this chapter, I have tried to resolve some challenges for the hard incompatibilist position regarding moral responsibility. In the absence of 'true' responsibility, there is still very strong case to be made to address a moral responsibility to people, but instead of it being based on a merit, like free will, I have shown that we can adhere a consequentialist justification of moral responsibility without being a compatibilist. Even more so, we appear to have good consequentialist reasons to be incompatibilist, rather than compatibilist, since it has us make some revisions in our moral responsibility that we have reason to believe are good. This can even be stretched to being an advantage of anti-libertarianism over libertarianism. I think these defenses for hard incompatibilism still stand, even though I have given a slightly different justification for this than my predecessors, in that I do not feel that all blame is necessarily wrong because it is undeserved, I would rather say that it is *non*-deserved.

In the next chapter, I will take on another topic important for my hard incompatibilist revisionist account, namely the role of our reactive attitudes. As my account is a bit more forgiving to by saying that blame is only non-deserved, I think I can make a more realistic case for revisionism there.

3. Moral responsibility and our attitudes.

Because our moral intuitions are thought to be (at least partly) based on moral emotions, such as anger, it appears then that they are natural. It might then also be the case that they are what mainly defines our ways of holding people responsible, and not our reason. Based on this idea, P.F. Strawson, in his benchmark work *Freedom and Resentment* (1962), tried to completely turn the determinist moral responsibility discussion upside down by suggesting that the reactive attitudes make for the legitimacy of our responsibility practices are and not the other way around. How we are supposed to blame each other does not depend on a whether we have a free will or not, or so he claims. The reactive attitudes, by which we hold each other responsible, are constitutive of personal interaction between autonomous adults.⁴⁹ They are appropriate for being genuine. If I take you to be morally responsible, I take you to be an adult and autonomous person. Thus opposingly, if I don't, I then take you to be non-autonomous and essentially an object of social policy, like a child or mentally ill person.⁵⁰ This is so because if I try to distance myself from my emotion based intuitions, like anger, that tell me that someone *is* responsible, I would take a strongly reasoned attitude instead of one that is emotionally involved.

The essential concepts added to the vocabulary of the anti-libertarian responsibility debate by Strawson are *reactive-* and *objective* attitudes. If we take an attitude towards someone or something, this means we have a certain 'mental stance' or position to it. For reactive attitudes, this means there has to be an *emotional* responsibility related attitude, essentially based on pre-reflective *reactive emotions*. I think there is a relevant distinction between these reactive emotions and reactive attitudes, yet somehow I never found this in the debate, except in passage by Sommers.⁵¹ Usually, people refer to both these concepts as being 'reactive attitudes' but I think this is wrong. The difference between these is that the reactive emotion is *primal*, and makes for the pre-reflective responsibility emotion. The reactive attitude can be taken on after reflecting on the emotions, or after not reflecting on them, but simply unreflectively keeping the emotional reaction. I think anger is a reactive emotion, although we can stay

⁴⁹ Strawson, "Freedom and Resentment", 7.

⁵⁰ Ibid, 9.

⁵¹ Tamler Sommers, "The Objective Attitude", *The Philosophical Quarterly* 57.228 (2007): 327

angry thus it might be a reactive attitude as well. Resentment on the other hand, is exclusively a reactive attitude. For us to resent someone, we have to first processed a few emotions to come to a conclusion that we strongly dislike this person. Anger, like fear, we feel immediately and activates for direct action (fight or flight). Still we can reflect on the anger and reason it away if it is wrong. So then, if we don't take on a reactive attitude, we take on what Strawson calls an "objective" attitude. By doing so, we are being rational in our reactions, and scrutinize all our emotions.⁵²

When accepting hard incompatibilism, one believes that that no-one is ever truly responsible for what one does. As the Basic Argument shows, one is not responsible for their ill will, so it appears that we ought to take a more objective attitude at all times. This incompatibilist position is called "abolitionist", as it opts for the abolition of reactive attitudes, like anger.⁵³ If we want to change our moral responsibility practices, it has to be changed at its roots. But this is exactly what Strawson and his followers challenge. Not only do they say that this is not what moral responsibility is about, but they also think that we *cannot* do this and that we *should not* do this.

Why they say we *cannot* do this is especially because of two reasons especially. First, it is so because that is simply how we are wired as human beings. As it is argued, our moral emotions, like anger, are simply a given and cannot be extirpated from our emotional repertoire. Secondly, it is thought that we cannot *want* to do this, because the world would become unlivable due to a lack of personal involvement. Even if we would somehow manage to take on an exclusively objective attitude, our world would become so extremely impersonal without these emotions, so that it would be 'shallow and bleak'.⁵⁴ Defenders of this position include Peter Strawson, Susan Wolf, and Seth Shabo. Why it is thought that we *should not*, even if we could, is because it is believed that our reactive attitudes are crucial for properly dealing out punishments. Our responsibility practices are dependent on these emotions, especially anger, to properly motivate us to apply deal out punishments. I will especially respond to Nichols' 2007 defense for this claim.

In this chapter, I will try to deal with this dual challenge on attitudes for the hard incompatibilist, as it is important for my revisionist idea of consequence based

⁵² Strawson, "Freedom and Resentment", 7-10.

⁵³ First coined in: Per-Erik Milam, "Reactive Attitudes and Personal Relationships", (2016).

⁵⁴ Susan Wolf, "The importance of free will." *Mind* 90.359 (1981): 386-405.

responsibility that this revisionism can be accepted. To do this, I will strongly follow the idea's and responses of my fellow abolitionists, and argue for the possibility, and desirability, of abolishing the reactive attitudes. To do this, I think it could be fruitful to consider the reactive emotions. While other hard incompatibilists desire to completely get rid of anger, I think I do need to go so far. I think some anger is natural, and to completely get rid of it is impossible, at least as a reactive emotion. But this does not mean we cannot get rid of the reactive attitudes that follow from it. Still, I do stand with the other incompatibilists regarding the goal of abolishing the reactive attitudes as we know them.

My approach in this chapter to most convincingly defend my position is by first showing that Strawson's move to base responsibility practices in the emotional reactive attitudes is unjustified. Second, I will deal with the claim that we *cannot* have abolitionism by showing that the anti-abolitionists are too culturally rigid in their views on emotions due to a bias, and discuss what approach might be most desirable. Then I will take on the challenge of the personal relations, and show that the non-reactive attitude that the hard incompatibilist has to take is not that impersonal and 'cold' as the compatibilist would have us believe. Finally, I deal with the claim that we should not outcome challenge, where I formulate a strong doubt on the usefulness of reactive *attitudes*, as compared to the usefulness of reactive *emotions*. I will do this by discussing the role of anger in moral motivation. I will conclude that those we view as being truly virtuous, take on a non-reactive, while still being motivated for a just cause.

3.1 Against the insulationist move.

The first challenge Strawson's account brings is that he bases his account of responsibility on something else than a free will, to him, it is rather based on a combination of genuine personal interaction and recognizing autonomy. It can be said that his compatibilism is essentially making an 'insulationist' move in the discussion of moral responsibility, by saying that whether we have a free will or not is essentially irrelevant for the reactive attitudes we take, or should take.⁵⁵ By doing so, he separates responsibility from free will, as if they were completely unrelated (hence the insulationist move). Therefore, this form of compatibilism denies the incompatibilist

⁵⁵ Nichols, "After Incompatibilism", 410.

step from the Basic Argument (step 3), and it would then not be wrong to blame someone that is not the ultimate source of one's actions. This would indeed be a very pleasant conclusion for compatibilists since anti-libertrianism could indeed be accepted without any moral implication.

Regrettably for the compatibilist, this insulationist move is currently regarded as mistaken by most philosophers, for especially two reasons that I will now shortly discuss. The first problem with the insulationist move is a cultural issue, in that it would disallow any fundamental criticism of any current moral practice. Strawson admits that our reactive attitudes are not only natural but also culturally defined. Our, biologically *and culturally* defined, reactive emotions would simply be justified, for being those that they are in their specific culture. It appears that we are not allowed to criticize the foundations of our responsibility practices, nor criticize them from outside.

“[I]nside the general structure or web of human attitudes and feelings of which I have been speaking, there is endless room for modification, redirection, criticism, and justification. But questions of justification are internal to the structure or relate to modifications internal to it. The existence of the general framework of attitudes itself is something we are given with the fact of human society. As a whole, it neither calls for nor permits, an external ‘rational’ justification.”⁵⁶

If our attitudes can only be thought to be appropriate or not from inside of their cultural framework, they would be completely impervious to external reasons. This does not match with our strongly held belief that criticism of a culture's practice of attitudes can very well be justified.⁵⁷ Strawson's compatibilism is thus a moral relativist position and is thus open for moral relativist criticisms, like that some cultural beliefs can be plainly *wrong*. But I will discuss this more extensively in the next subsection.

The second argument against the insulationist move can be aptly explained by the following murder trial example as described by Robert Kane:

⁵⁶ Peter Strawson, “Freedom & Resentment”, (1962), p. 23

⁵⁷ Laura W. Ekstrom, *Free Will: A Philosophical Study*, (Boulder, CO: Westview Press, 2000) 148-149.

“[a] young man who had raped and murdered a sixteen-year-old girl [...] My initial thoughts of the young man were filled with anger and resentment. But as I listened daily to the testimony of how he came to have the mean character and perverse motives he did have a sordid story of parental neglect, child abuse, bad role models, and so on some of my resentment toward him decreased and was directed toward other persons who abused and influenced him. [...] In such manner, the changes in reactive attitudes [...] are related to beliefs about ultimate responsibility.”⁵⁸

I think this example vividly shows the incompatibilist response to the insulationist move: the reactive attitude shifts when the perceived source of the evil shifts. It appears that our reactive attitudes *are*, in fact, strongly dependent on the actual source.

How we hold one responsible is not justified merely by our current culture’s moral emotions norms, nor independent of the actual source. For these two reasons, I will content that the insulationist move does not work and that compatibilists thus have to give up on their belief that we can blame people based simply on their reactive attitudes. Additional justifications need to be given why we would be right to take on these attitudes towards others.

3.2 The impossibility challenge

I agree with Strawson’s claim that how we as humans come to initially think of one as responsible is based on our normative dispositions, or: reactive emotions. On the other hand, these reactive emotions are still *unreflective*, and can very well make for punishments that have in the long run bad consequences if we act on them unreflectively. Because of this,⁵⁹ abolitionists plead to get rid of the reactive attitudes, like anger, resentment, and guilt, because they all make for unfair blame. Instead, some opt that we ought to replace the reactive emotions with *non-reactive emotions*, like

⁵⁸ As a side note on this example by Kane: Kane’s shifting of resentment from the murderer to his parents (and other abusers) is not justified according to hard incompatibilism either, because his parents are just as not responsible for who they happened to be. We have to expect that they have had problematic childhoods that led them to have this immoral character as well, or they just have bad genes. Either way, we would already enter the regression of responsibility as discussed in the Basic Argument.

Kane, “Significance of Free Will”, 84.

⁵⁹ As well as other reasons, like the fairness of *any kind* of blame. However, as I have discussed, I am less keen to support the fairness claim.

sadness and regret, for the fact that these are essentially non-blaming, and have us take on a more objective attitude meaning that we need reflection before moral judgment.⁶⁰

It thus appears we have debatable whether we want them, and I will try to settle this debate in section 4.3. But first, I will have to assess to what extent this question is even worth asking. Even if we would say that we do not want these attitudes, we might be stuck with them anyway. The impossibility challenge is essentially twofold and contains both an empirical claim and a normative claim. The empirical claim is that it is (nearly) impossible to change our emotional repertoire, and the normative claim is that a livable society is impossible where one constantly takes on objective attitudes. In the following section, I will try to deal with both of these claims to make a sufficiently strong case for the revisionist, as he would not be prematurely disallowed to have his abolitionist beliefs.

3.2.1 'Natural' emotions

Strawson does not doubt whether we should take on reactive attitudes or not. The question does not even make sense to him: "it is useless to ask whether it would not be rational for us to do what is not in our nature to (be able to) do."⁶¹ However, we have reason to doubt whether this is truly the case, most specifically because of the fact that there is great cultural variation in our moral emotions. Which, as I have already mentioned, Strawson does recognize. In the following subsection, I will show what kinds of implications this role culture has to the impossibility challenge of 'naturalness'.

First of all, I think it is inconsistent by Strawson to claim that we cannot do other than act on our moral emotions on responsibility, while at the same time recognizing the role of cultural influences. This is so because cultural norms can *change*, and often rightly so. We can think of many norms that we might think have been in one way or another *understandable* for culture to have at that place at that time, but that does not mean that they were ever *good*. Take for example the long history of slavery. As mentioned in the previous section, we do generally think that certain cultural norms that might very well be intuitive to people from within that culture need to be open for external criticism. Yet, some kinds of cultural beliefs and practices are harder to change

⁶⁰ Sommers, "Objective Attitudes", 328-340; Pereboom, "Free Will, Agency", 146-151.

⁶¹ Strawson, "Freedom and Resentment", 14.

than others. These beliefs are especially those that concern our emotions. For example, if we would have shown Aristoteles our present-day evidence that non-Greek people had just as large brains as Greek people and are just as intelligent, then he might have revised his belief on whether they would be okay to be held as slaves.⁶² Instead, such fact based approach is probably much less fruitful in this situation considering a emotionally laden belief by another great philosopher: Immanuel Kant had expressed a significantly clear disgust when discussing homosexuality and masturbation. He described it as "*Crimina carne contra natura*" ("crimes from the flesh against nature"),⁶³ and put these phenomena in the same row with bestiality. Being embedded in 18th-century Königsburg, we could say that it is understandable for him to have such belief and even possibly even his emotional reaction, but not good. Now, I would have liked to make the point that in present Königsburg, the standard emotional reactions to this is very much different, by being 21st century Germany,⁶⁴ and that therefore our emotional dispositions within a culture can change. It thus might take a few centuries for a cultures norms to change appropriately, even if there is a good reason to not have these emotional reactions against homosexuality and masturbation all this time.

Why it is hard to change an emotionally laden belief is because it becomes extremely *intuitive*; it seems obviously true. People tend to perceive them as something that is *more* than their subjective experience of it, but rather an independent *truth* about the matter.⁶⁵ A cultural, emotionally laden, normative belief can become so vivid that people have trouble thinking of even the possibility of it not being true. This bias of thinking an often culturally based emotion is actually the 'only way' instead of 'your way' is called 'ratchet effect'.⁶⁶ Another historical example of this ratchet effect is by Herodotus on an interesting and emotional event at the court of the Persian king Darius [Histories 3.38]:

⁶² Wedgwood, "Moral Disagreement", (2014), 25.

⁶³ Immanuel Kant, 'moral Collin's lecture notes. Lectures on ethics. The Cambridge edition of the works of Immanuel Kant. 160-1

⁶⁴ The intention was to make a this comparison, but then I learned that the old Prussian city of Konigsberg is better known as 'Kalingrad' nowadays. It is now inhabited by Russians and in Russian territory after the Second World War. Homosexuality is much less accepted in Russia there than it is in Germany, thus making this specific case a less apt example for my point than I intended it to be, but I presume that the message is clear regardless.

⁶⁵ Wedgwood, "Moral Disagreement", 32.

⁶⁶ Owen Flanagan, *The Geography of Morals*, 184.

“When he was king of Persia, he summoned the Greeks who happened to be present at his court, and asked them what they would take to eat the dead bodies of their fathers. They replied that they would not do it for any money in the world. Later, in the presence of the Greeks, and through an interpreter, so that they could understand what was said, he asked some Indians of the tribe called Callatiae, who do in fact eat their parents' dead bodies, what they would take to burn them. They uttered a cry of horror and forbade him to mention such a dreadful thing. One can see by this what custom can do.”

Both this example and the example of Kant show how impossible certain practices can seem, while if they merely go against our cultural emotional intuitions. I think this ratchet effect shows that we ought to be careful to not prematurely dismiss certain practices if they are not in line with our emotions and intuitions. If we are thus to accept that we think some cultures have rightfully changed their emotional dispositions, in light of some morally relevant knowledge concerning their ‘original’ practices, then why shouldn’t we, in light of relevant information?

It is clearly not the case that our current emotional phenomenology is exactly like that of the first men, so it is not the natural one.⁶⁷ We cannot arbitrarily assume that our western set of emotional reactions is more true than others, due to the bias of the ratchet effect. So where thus this leave us? Although we are culturally embedded, and it might be impossible to take a truly evaluative stance regarding ourselves including our emotions and our culture,⁶⁸ it might not be impossible to imagine a cultural emotional repertoire that is more in line with a hard incompatibilist view. Or just one that is more morally desirable overall.

When looking for a new emotional repertoire, I would like to make an honorable mention to the Stoics, as well as a variety of branches of Hinduism and Buddhism. These groups have at least in some part of their beliefs a psychological goal that is in line with the abolitionist: to banish anger.⁶⁹ Although admirable, I strongly doubt the practicality

⁶⁷ Flanagan describes, following William James, that the different words we use to describe our emotions are not only important for the noumenal view one has for their emotions, but that it is so pervasive that it makes up our *phenomenal* experience as well. Owen Flanagan, *The Geography of Morals*, 180.

⁶⁸ Thomas Nagel, “Moral Luck”, *Proceedings of the Aristotelian Society*, 50 (1976), 152.

⁶⁹ Flanagan, *The Geography of Morals*, 179.

of this for my thesis, for one reason specifically: in neither present-day Buddhist or Hinduist cultures there appears to be a collective internalization of these emotional norms, with the exception of a relatively small esoteric elite who achieved this through years of rigorous training. My goal is not to argue for universal *asceticism*, for this is overly demanding, but to argue for realistic emotional norms we might reasonably take on. I argue that due to their over-demandingness by aiming to get rid of the reactive emotion of anger, this is where I argue they wash up. But as I propose a less extreme version of abolitionism, this need not affect my theory as much as theirs.

What I think goes wrong here, is that the Buddhists and the Stoics try get rid of the *reactive emotions*, instead of the reactive attitudes, just as the other abolitionists seem to try. This is their mistake. Anger is in us. But how we deal with it is what matters. In the Ifaluk tribe of the Caroline Islands, for example, pre-reflective anger is especially expressed in a loss of appetite, and it is almost never hostile,⁷⁰ while in western cultures anger is often described to elicit a desire to retaliate or in any way get back at the person who we feel mistreated us.⁷¹ Since we have good reason to doubt the appropriateness of our anger, it would be an improvement to have as an pre-reflective anger be one that is by default quite harmless rather than harmful. Ifaluk anger is much less *reactively expressed* than western anger, so don't we then have a reason to change our default reactive attitude to that of the Ifaluk? I think we do.

However, I think there is a realistic worry about some crucial differences between the Ifaluk society and ours that might disallow us to take on their cultural emotional norms. The Ifaluk are a small society, where people are very much engaged in the group, as opposed of our western globalized capitalistic society in which we meet new people everyday and we don't need to truly invest in all of them. The capitalist, or neoliberal, ideal is rather that if we all just pursue our own interests things end up best for all. A western expression of anger in which we immediately get angry and try to stake our claim makes much more sense when understood within such ideals. But as I have described this of other emotional norms before: I think this is *understandable*, not morally correct. I am curious how much of a barrier this capitalist society is for our

⁷⁰ Flanagan, *The Geography of Morals*, 189-190.

⁷¹ Nichols, "Hard Incompatibilism", 413,

cultural emotional revisions. I think this is an interesting question to explore in future research, as this is currently lacking.

3.2.2 Impossible to live in?

The second impossibility challenge makes a claim on the crucial role of reactive attitudes on our *personal relationships*. This is thus specifically a worry on the incompatibilist's abolitionist goal, which suggests we ought to get rid of our reactive attitudes entirely and instead take on exclusively an objective attitude. It is feared that taking on these objective attitudes would preclude genuine personal relationships.⁷² This poses a significant problem for revisionism by means of abolitionism, as the world would be one we could not live in. Susan Wolf describes a such world as one that is "A world of human isolation so cold and dreary that any but the most cynical must shudder at the idea of it".⁷³

If Wolf's description is correct, the abolitionist position becomes impossible to hold. It might be the case that the revisionists ought to stop taking an abolitionist position, and merely focus on other institutes of our moral practices that need revision in light of indeterminism. However, in this subsection, I will try to shed a reasonable doubt on Wolf's description, as it is not at all clear that these reactive attitudes are indeed as important for personal relationships as Wolf claims they are. But why would they be true? Wolf does not give us any good reason to think that they are. But Seth Shabo does.⁷⁴ This argument regarding the crucial role of reactive attitudes in personal relationships, specifically *mature friendship* and *reciprocal love*, can be called the 'Personal Relations Argument'. It can be deductively formulated as follows:

1. If one is not disposed at all to take reactive attitudes towards an agent, then one will not be disposed to take her actions or attitudes personally.
2. If one is not disposed to take an agent's actions or attitudes personally, then one cannot have a personal relationship with her.

⁷² Strawson, "Freedom and Resentment", 9.

⁷³ Wolf, "Importance of Free Will", 392. For a more complete description of this phenomenon see Wolf's own passage. The vivid description there is quoted by both Sommers (2007) and Shabo (2012). It is the basis for their discussion on the attitudes.

⁷⁴ Seth Shabo, "Incompatibilism and Personal Relationships: Another Look at Strawson's Objective attitude." *Australasian Journal of Philosophy* 90.1 (2012): 131-147.

3. Therefore, if one is not disposed at all to take reactive attitudes toward an agent, then one cannot have a personal relationship with her.⁷⁵

There is an intuitive pull to this argument. We tend to get emotional as a reaction to when something does something kind or unkind to us personally. However, after closer inspection a variety of quite problematic claims would be correct according to this argument. Milam (2016) points out that the following quite problematic propositions are implied by this argument from Shabo:⁷⁶

1. If you cease to be friends with someone, your disposition to hold feelings of resentment to that person becomes less strong (p. 108)
2. Friendships in which the people involved sometimes resent each other are more healthy and valuable than those in which they don't, *all other things being equal*. (p. 108)
3. A world where people would at all times be kind to each other would be worse than where people are mean to each other from time to time. (p. 108)
4. People cannot take anything a non-agent (child, mentally ill, animal) says or does personal, as they are not autonomous beings. (p. 109)
5. Non-reactive emotions like sadness, disappointment, or embarrassment cannot make for something to be experienced as personal. (p. 110)

These problematic propositions are what Milam describes 'incompatibility problems'. They are meant to show that the line that Shabo draws between reactive attitudes and non-reactive attitudes regarding personal relationships is clearly different from how we see it. Taking things personal is rather related to something as being targeted at us specifically, this does not require a reactive emotion or attitude. For example, point 4 is derived from an example where a toddler daughter says something hurtful to her father

⁷⁵ I take this formulation of Shabo's argument from Per-Erik Milam "Reactive attitudes and Personal Relationships", *Canadian Journal of Philosophy*, 46, no 1, (2016): 106-107.

⁷⁶ Obviously these propositions are taken somewhat out of their context, but I do think each of these are apt *reductios* of Shabo's argument. For a more in-depth analysis and explanation of these problems with the Personal Relations Argument, please see Milam's original essay, as I cannot in length repeat the exact context of all these propositions.

about one of his personal insecurity she knows of; point 5 relates to an example where someone was denied a job due to a form of racism, where the solicitant feels *sad* instead of *angry*. We that in both cases they still can take it personal.

If these problems indeed show that Shabo's account on personal relationships is not as strong, we are free to accept abolitionism again without worrying that our ability to take things personally, and thus our personal relationships, would be made impossible. Our world might after all not be so cold and bleak when taking on an objective attitude. I think this also shows that 'objective' as an adjective to these non-reactive attitudes is inaccurate.

Another example might make it even more clear that non-reactive attitudes are not really what we would call objective: A good friend, Pablo, mistreats me. I don't like that at all, and I take it personally since I know that he would not have done that to someone else. The harm was directed at *me* specifically. For this reason, I can be saddened and disappointed because what I thought was a valuable mutual friendship is not really what I thought it was. I am in grief, for I have always valued my friend's company, but now I feel I cannot expect from him any longer the pleasant interaction and loyalty that make a friendship valuable. If this is description is correct, it appears that one can still surely *value* things in personal relationships without taking on reactive attitudes. Valuing is something that is generally thought of as almost purely *subjective*,⁷⁷ and subjective is typically seen as the opposite of objective. If thus an attitude allows us to value things, it is thus still at least partly subjective; this attitude could not be thought of anymore as being objective. I thus say the term 'non-reactive' makes for a more correct and less confusing terminology and raises less fear of potential bleakness.⁷⁸

Finally, there is another interesting feature of taking a hard incompatibilist non-reactive attitude, although it is not too personal. Still, it is the case that they allow for a way to elicit *more* emotional involvement into our judgments, especially towards the *wrongdoers*, whose position we sometimes tend to forget. When taking on a non-reactive attitude, we do not form our judgments unreflectively as when taking on reactive attitudes but take instead reflect and take a more reasoned stance on the situation in order to find out what reaction is appropriate. When doing such reflection,

⁷⁷ Max Wever, "Die" Objektivität" sozialwissenschaftlicher und sozialpolitischer Erkenntnis." *Archiv für sozialwissenschaft und sozialpolitik*, 19(1), (1904): 22-87.

⁷⁸ This namecalling has been identified by Sommers, "The Objective Attitude", 325.

we can also consider the Basic Argument. Doing so, we come to realize that the wrongdoer's nasty actions are ultimately not his own. Causes for one's problematic character features are typically either anti-social genes or social factors. People of bad moral character often tend to get a lot of conflict in their lives. Thus generally, hard incompatibilism suggests that criminals are essentially people who were very unlucky social-economically or genetically, and therefore one is more likely to imagine him as a *victim* as well, someone who has tragically been stricken by some terrible luck in life. Therefore, accepting and properly understanding hard incompatibilism causes a *reasoned empathy*,⁷⁹ or rather: compassion. We can see how this compassion works when thinking of the murder trial example by Robert Kane: As our resentment shifts and eventually dissolves to events, and we become more compassionate to the murderer. This way, it clearly becomes much easier to be constructive towards criminals and have consequentialist punishments, that are typically much less harsh. This way, our emotions can come to be in line with our consequentialist reason, making it much easier to think of less retributive punishment as good, as we take more interest in the criminals' position, and yes, this does not only work for considering criminals, but it allows us to be more compassioned to all.

In this section, I discussed that cultural moral emotions are sometimes wrong, but also that they can, and do, change. This change might be difficult, as often people cannot imagine taking on another view if theirs is so much emotionally laden. Because of this, emotional reforms in a culture can take quite a while. But this does not mean we can achieve abolitionism. I think that the reactive emotions remain, as long as we do not actively extirpate by means of ascetic effort, like the Stoics or Buddhist monks. Thus, it might be more fruitful to change their *pre-reflective expression*. The cultural norm of the reactive emotion of anger of the Ifaluk tribe is potentially a hope giving example, as it would make it much less problematic for our emotions to be thus expressed if we fail to attain our reflective non-reactive attitude. The Ifaluk reactive attitude would be better than our a more retributive western one if we slip up and accidentally blame but still, the attainability of this needs more exploring as it might be difficult to fit in our capitalist society.

Additionally, it is quite clear that these reactive attitudes are not as crucial to our

⁷⁹ Nichols, "After Incompatibilism", 415.

personal interaction people tend to think them to be. Likely this idea has come into this world by the unfortunate name that Strawson gave to the attitudes that are not reactive, namely 'objective'.⁸⁰ The non-reactive attitude might even add more emotional involvement: as it can help us achieve more *compassion*. Through incompatibilist reason, we can reason an attitude by means of which we take more consideration for others' beliefs. I think this shows clearly the moral potential of changing our attitudes, as well as our cultural expression of the reactive emotions. I think there is a lot of potential here to improve our lives. Yet this final claim still has to be questioned in the following subsection concerning outcomes.

3.3 The Outcome challenge

If indeed the world without reactive attitudes would not be overly bleak and barren, it might be one we could very well live in. It might even be better than living in our current libertarian society. Even though there are some good reasons to get rid of our reactive attitudes, there might also be other quite stringent reasons for us to hold on these attitudes over the abolishment of them. The compatibilist claim that I will discuss in this section is that abolishing our reactive attitudes would undermine a well-functioning morality, since these specific emotions are crucial for our system of holding each other responsible effectively. Specifically, this comes down to the worry that our non-reactive emotions fail to cover for the uses of the useful moral emotions that we are abolishing, like anger.

Anger could be thought of as being *the* moral emotion that largely defines our morality.⁸¹ The most common cause for someone to experience anger is experienced injustice.⁸² If one experiences anger, one is generally *activated*. Anger motivates us to revolt against injustice,⁸³ and thus if someone harms us in a way we think is not okay, we tend to draw a line by getting angry thereby showing her that this was not something that she was allowed to do. But we already know that our western experience of anger is often quite different from other cultures and as we have learned from the ratchet effect, we ought not to confuse our western way with the natural way.

⁸⁰ Sommers, "Objective Attitude, 325.

⁸¹ Nichols, "After Incompatibilism", 413.

⁸² Ibid, 413.

⁸³ Ibid, 416-421

Incompatibilists tend to bring forth non-reactive emotions, like sadness, as being apt replacements for the reactive emotions.⁸⁴ Sadness expresses a regret that something happened, as mourning. Sadness, like anger, expresses a value judgment, simplified: “I do not like X”, X being the cause of your emotion. Compatibilists, however, argue that sadness is too much of an emotion that leads to *inaction*. They are more likely to get depressed or demoralized by the harm, which is clearly not what we think angry people tend to do. Angry people stand up and make a ruckus. Additionally, Nichols argues that people tend to base their punishments on their anger emotions, by referring to an economic game study. In this study, people can cooperate, freeride, or punish.⁸⁵ Those who chose to free ride were frequently punished by the others, after which they would more often cooperate. He argues that, because those who wanted to cooperate got angry, this incentivized them to punish the free riders. If not for this anger induced punishment, it would be unlikely that the free riders would have been properly addressed. To Nichols, this proves that we need our anger emotion to properly address injustices.

Incompatibilists have offered some relevant rebuttals on this, specifically Pereboom (2014). For one, Nichols bases the idea of inaction caused by sadness on *infant research*.⁸⁶ Infant emotions likely have very different experience and functions than that in an adult, such as the goal to elicit a caring empathy by its parents. Instead, Pereboom argues that Nichols account lacks recognition of rational *moral resolve*.⁸⁷ Some people have strong moral motivation, regardless of their reactive emotions.

Another response that I think is worth considering is that anger is in fact not as much of a reactive attitude as it is as *reactive emotion*. However, but you *can* reflect on this emotion before you act on it, in most cases. Before you have done any reflection, you feel the anger. Opposingly, reactive attitudes like resentment are emotions that have been given much more time. As in the murder example, we can take the resentment

⁸⁴ Pereboom, “Living Without”, 146-151.

⁸⁵ Nichols, “Afterincompatibilism, 418. Referring to Fehr & Fischbacher.

⁸⁶ At page. 420 of “After incompatibilism” (2007) Nichols disqualifies alternative emotions (like sadness) as being apt to motivate one to stand up for justice, by saying that these other emotions are ineffective in doing so, he quotes Lazarus (1991, p. 251): “In sadness there seems to be no clear action tendency-except *inaction* or withdrawal into oneself”. However, it is argued by Pereboom (2014, p. 148) that this study by Lazarus is not a very convincing account on this, as it is a study that is specifically about *infants*. Clearly, there is a significant difference in adult and infant emotion based motivation.

⁸⁷ Pereboom, “Free Will, Agency”, 149.

away using reason and thus stop having this attitude. This is much harder for the reactive emotions like anger, it would require an ascetic kind of abolishment. Take for example Sommers' explanation of this: If my TV got stolen, I will likely *initially* feel anger towards the thief, next to my non-reactive sadness, but after having thought of the thief as being raised poorly, among other things, I cannot really come to resent him.⁸⁸

When talking about moral motivation without taking on reactive attitudes, one can, for example, look at Martin Luther King, Mahatma Ghandi and Nelson Mandela. All of these men are looked at as great leaders that have revolted against perceived injustice without any apparent expression of anger or resentment at all, but with peaceful protest.⁸⁹ Of course, I cannot just say: "look at these men, it is simple, just do as they do". This is clearly much easier said than done. These men's virtue is of legendary status. But what is also clear is that part of what we admire in them is *exactly their refraining from taking reactive attitudes*. Being *virtuous*, in these cases, is to revolt in a non-reactive way.

Perhaps against this account the compatibilist might respond: "Even though they acted seemingly rational and not angry, in fact, they must have been angry first, in order to be motivated to act against injustice at all", in other words, what Pereboom describes as being 'moral resolve' is actually *still* a form of anger. I think this might be the case. But this is not anger as being a reactive attitude, it would only be anger as a reactive emotion. Or, perhaps, a variation of anger in its wider psychological profile. What ever kind of anger it might have been. These virtuous men, they did not *act* angry, or at least not with any *rage* or *fury*. We think that their approach was as reasonable as their cause. They must thus have strongly reflected on their emotion, until the attitude they have taken, was one without retributive desire. Or, alternatively of course, their initial anger expression was already minimally reactive. A third explanation might be that they had more empathy by nature.

Therefore, I think we can quite safely say that they took on a non-reactive attitude, one way or another. But it is less sure that they have achieved the ultimate goal for the abolitionist: the *total* extirpation of reactive emotions. But if even *they* then supposedly failed for this theory, I think the abolitionist is clearly overly demanding.

⁸⁸ Sommers, "Objective Attitude", 328.

⁸⁹ Pereboom, "Living Without, Agency", 149.

3.4 Conclusion

I think my account in this chapter ought to give a good view of what it would mean to take on a reactive attitude and what it means to not do so. First of all, it might be useful to distinguish between the pre-reflective emotion, the reactive or non-reactive emotion and the attitude one ends up with. These emotional reactions differ per culture, they are elicited by different things and expressed in different ways, and some of these we have reason to think are morally preferable than others. Some cultural revision might thus be required, although this might be difficult considering our western capitalist embedding, but not necessarily impossible and undesirable. For the time being, it is best to hold on to the ideal to try not to act on them unreflectively. Rather as to not necessarily root out the reactive emotions, but do not allow for reactive attitudes to be formed, by taking a non-reactive attitude. I think distinction is something that needs to be taken a more critical look at in the future, as most philosophers tend refer to both of these concepts by the term 'reactive attitude'. Furthermore, I think my theory can be thought of as being more allowing in these practices, and therefore is less strongly abolitionist than my predecessors, which I think it is much more realistic for achieving a desirable moral emotion revision.

I think this all shows that consequentialist based responsibility practices can take on the final criticism, our emotional reactions. As reactive attitudes are not as important in any of those ways the compatibilists have tried to defend, then we have good reason to change our emotional reactions in line with hard incompatibilism.

Conclusion & Final remarks

If we ask: “what’s at stake at the free will debate?” we tend to think of what we have to lose. But in this thesis, I have showed that it is highly unlikely we have much to lose. In fact, how I sketched and defended hard incompatibilism, it appears we mainly have things to *gain* when we let go of our free will belief, as certain of our problematic moral practices, like punishing harshly, tend to lose their most prominent grounding.

In this thesis, I have attempted to make up the score in the anti-libertarian discussion. My goal was to make a strong and reasonable account for hard incompatibilist revisionism, by defending it against a variety of challenges. Based on the assumption that we can still very well be reasonable beings in an (pseudo)determinist anti-libertarian world, the first step in achieving my goal was by salvaging the most important moral practice that is changed by hard incompatibilism, namely moral responsibility. I discussed a consequentialist based account of responsibility that is very well in line with the hard incompatibilist position, without having to accept a weaker metaethical position. Furthermore, the position is not much more likely to obtain morally problematic results like opting for ‘weird’ punishments.

As this envisioned revised responsibility practice allows us, other than any merit-based practice, to opt *exclusively* for desirable consequences, it is to be expected that the consequences are better. For good consequences, revisionism is necessary in light of anti-libertarianism, especially in the justice system, meaning that the consequentialist responsibility account would have us accept leans towards a incompatibilist moral change. Additionally, the only justification for not achieving optimal consequences appears to lose its moral fundament, thereby making the moral changes that happen for good consequences easier to accept in an anti-libertarian view, where excessively harsh punishments are not deserved. Crucial for achieving this, is that it is necessary that we need to *reflect well* on our responsibility related intuitions, since they are much more in line with the libertarian or compatibilist merit-based belief that people *are responsible*. By this need of having to be systematically skeptical of these reactive emotions, we might have reason to abolish them overall, or at least significantly change how they are expressed. I think the latter of these is more realistic.

I thus explored the option of hard incompatibilist emotion revisionism, namely abolitionism. Since Peter Strawson’s isolationism was wrong, it appears that it *does*

matter people are the source of their action. We thus have more reason to believe some of our emotionally laden beliefs are mistaken, and although we know it is difficult to change them, we know that at least part of them is cultural, like how they are expressed and when. Because we know that this is the case, it is possible to change them, as we have before in the past with several of our problematic emotionally laden beliefs. Furthermore, the alternative attitude that one takes, is rather non-reactive than objective. We do still have our non-reactive emotions, and value things. We can gain more compassion when reasoning in a determinist fashion, which has obvious moral benefits. Even if one reactive emotion might be useful for moral resolve, the actually desirable moral reaction does not display any anger anymore. Acting against evil, without anger or hate, is what we tend to view as morally virtuous.

Therefore, I think I have made a good defense for incompatibilism by taking a way multiple commonly understood fundamentals based on which the compatibilists justify the reactive attitudes. But if we can indeed not abolish the reactive emotion of anger, which I think we have good reasons to believe is the case, we could achieve a morally justified cultural change in *how* we express them, unreflectively. It is then still best to keep a reflective stance to these attitudes at all times as this might turn out to be a stance that is, at least by the looks of it, non-reactive. Since being reflective on our emotions at *all* times is likely an somewhat ascetic task in itself, it is best to have already changed the western pre-reflective emotion expression when failing to reflect. I have argued, for the attainability of this change I think that we need to take a good look at the role of capitalism in the make up of our western emotional cultural norms. I thus think the abolitionist position is empirically not sufficiently well supported, but the above described emotional revisions also work for my position, for my incompatibilist is a bit *softer*; because blame is always rather non-fair, and thus prima facie wrong, instead of *unfair*.

If my argumentation is correct, hard incompatibilism provides us good consequentialist reasons for certain changes in our moral responsibility practices, including our justice system by allowing for less harmful and more effective punishments, and it shows how to change our emotions, and how we handle them, for the better. I think that compatibilists therefore have to give up their compatibilist beliefs, since there is nothing to lose when letting go of our free will belief. Instead, I

have shown that there is a lot to *gain*. I thus invite you to hope with me that philosophers or other researchers find compelling (quasi-)determinist evidence against libertarianism, for it can, when well executed, result in desirable changes.

References

Aristotle, *The Nicomachean Ethics*, Book III, Translated by W.D. Ross. (1928), updated version 2015, taken from

<https://ebooks.adelaide.edu.au/a/aristotle/nicomachean/book3.html> 17-6-2018.

Dennett, Daniel Clement. *Elbow Room: The Varieties of Free Will Worth Wanting*. Boston: MIT Press, 1984.

Ekstrom, Laura Waddell. *Free Will: A Philosophical Study*, (Boulder, CO: Westview Press, 2000) 148-149.

Fischer, John Martin and Ravizza, Mark. 1998. *Responsibility and Control: A Theory of Moral Responsibility*, New York: Cambridge University Press.

Flack, Jessica & De Waal, Frans, “Any Animal Whatever”, *Journal of Consciousness Studies*, 7, no. 1-2 (2000).

Flanagan, Owen, *The Geography of Morals: Varieties of Moral Possibility*. Oxford: Oxford University Press, 2016.

Herodotus. [Histories 3.38]

Kane, Robert. *The Significance of Free Will*. Oxford: Oxford University Press, 1998.

Kant, Immanuel. *Moral Collin's lecture notes. Lectures on ethics*. The Cambridge edition of the works of Immanuel Kant. 160-1.

Kant, Immanuel. *Grounding for the Metaphysics of Morals*. (1993) [1785]. 36. 4:429.

Kleiman, Mark Albert Robert. *When Brute Force Fails*. Princeton: Princeton University Press, 2009.

Levy, Neil. *Addiction and Self-control: Perspectives from Philosophy, Psychology, and Neuroscience*. Oxford: Oxford University Press, 2013.

Levy, Neil. "Less Blame, Less Crime? The Practical Implications of Moral Responsibility Skepticism." *Journal of Practical Ethics* 3 (2) (2015): 1-17.

Lyons, Christopher & Pettit, Becky. 2011. Compounded Disadvantage: Race, Incarceration, and Wage Growth. *Social Problems* 58: 257-280.

Milam, Per-Erik. "Reactive Attitudes and Personal Relationships." *Canadian Journal of Philosophy* 46.1 (2016): 102-122.

Milam, Per-Erik. "In Defense of Non-reactive Attitudes." *Philosophical Explorations* 20.3 (2017): 294-307.

Nagel, Thomas. "Moral Luck," *Proceedings of the Aristotelian Society*, 50 (1976), 137–55.

Nietzsche, Friedrich, *Beyond Good and Evil*, (1886).

Pereboom, Derk. *Living without Free Will*. Cambridge: Cambridge University Press, 2001.

Pereboom, Derk. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press, 2014.

Sehon, Scott. *Free Will and Action Explanation: A Non-causal, Compatibilist Account*. Oxford: Oxford University Press, 2016.

Shabo, Seth. "Incompatibilism and Personal Relationships: Another look at Strawson's Objective Attitude." *Australasian Journal of Philosophy* 90.1 (2012): 131-147.

Smilansky, Saul. "The ethical advantages of hard determinism." *Philosophy and Phenomenological Research* 54.2 (1994): 355-363.

Smilansky, Saul. d Determinism and Punishment: A Practical Reductio." *Law and Philosophy* 30 (2011): 353-367.

Sommers, Tamler. "The objective attitude." *The Philosophical Quarterly* 57.228 (2007): 321-341.

Strawson, Galen. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75.1-2 (1994): 5-24.

Strawson, Peter F. "Freedom and Resentment," *Proceedings of the British Academy*, 48 (1962), 1-25.

Street, Sharon. "A Darwinian dilemma for realist theories of value." *Philosophical Studies* 127.1 (2006): 109-166.

Taylor, Richard, "Determinism and the Theory of Agency," *Determinism and Freedom*, New York University Press, (1959), 211-18.

Thomson, Judith Jarvis. "Killing, letting die, and the trolley problem." *The Monist* 59.2 (1976): 204-217.

Weber, Max. Die "Objektivität" sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Archiv für sozialwissenschaft und sozialpolitik*, 19(1), (1904): 22-87.

Wedgwood, Ralph. "Moral Disagreement among Philosophers", in, *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Michael Bergmann & Patrick Kain (eds.) (Oxford: Oxford University Press, 2014): 23-39.

Wolf, Susan. *Freedom Within Reason*. Oxford: Oxford University Press, 1993.

Wolf, Susan. "The Importance of Free Will." *Mind*, 90, 359. Oxford University Press, (1981): 386-405.