The Rationality of Anger

Laura Luz Sousa Oliveira e Silva

UCL

A thesis submitted for the degree of

Doctor of Philosophy

Department of Philosophy

University College London

July 5th, 2019

I, Laura Luz Sousa Oliveira e Silva, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

_____

Abstract

Anger has a bad reputation. The Stoics have famously advocated against anger (Nussbaum 2015; Bommarito 2017), while philosophers interested in oppression, particularly feminist philosophers, have on the other hand, highlighted the crucial motivational, political and epistemic value of anger in fighting against the oppressive status quo (Lorde 1981; Frye 1983; hooks 1995; Yancy 2008). I argue that the points of disagreement that emerge from the debate on anger hinge on its nature and (ir)rationality. Disagreement hinges on whether anger is instrumentally beneficial to the angry agent, in terms of its psychological effects and behavioural consequences, on whether anger is of epistemic value, and on how to construe the nature of anger itself.

A foundational approach to the debate, one that makes the emotion of anger central, will reveal that anger is a rational phenomenon that can play the positive roles feminist philosophers have afforded it. I provide such a foundational account of anger. The project has four complimentary desiderata; to provide an account of anger's nature, an account of anger as instrumentally rational for oppressed agents, an account of anger as reason-responsive, and finally an account of anger's positive epistemic role. I work within two methodological constraints: not to abstract away from how social reality is structured, and to be informed by the most relevant empirical work in the brain and behavioural sciences.

In meeting my four desiderata, I provide a currently lacking account of anger as rational. This provides the foundation for its constructive political roles. Four main findings emerge: that the dominant conception of anger should be revised, that anger is an effective and often constructive way of confronting social injustice, that anger, as an emotion, enjoys a sui generis relation to its reasons, and that this in turn allows anger to play distinctive and crucial epistemic roles.

Impact Statement

This project takes place at the intersection of fields that are, for the most part, undergone in isolation from one another. Building bridges across them is, I think, reciprocally beneficial for research in these fields. First, feminist philosophy and contemporary philosophy of emotion have not interacted with each other nearly as much as one might expect. Despite feminist philosophers having been instrumental in challenging the dichotomy between reason and emotion, a dichotomy which contemporary philosophers of emotion have long dismissed, the insights of these respective literatures have not typically been brought to bear on one another. Accounts of the nature and epistemology of emotions in the philosophy of emotion have, typically, abstracted away from facts of social reality, while feminist philosophers, on the other hand, have not often sought to provide specific accounts of the metaphysics and epistemology of the emotional phenomena that feature in their theorizing. This thesis takes steps to bridge these gaps by providing a foundational account of anger that allows it to play the sorts of roles feminist philosophers have afforded it.

This project makes use of empirical work from the brain and behavioural sciences. In so doing, it represents an empirically informed contribution to both contemporary philosophy of emotion, and feminist moral psychology. Beyond philosophy, chapters 2 and 3 are likely to be of interest to affective scientists as well, as they involve theoretical framing a range of experimental work on anger. They offer a non-reductive, but empirically sound, account of anger, one that sits well with experimental psychology, philosophy of mind, and feminist moral psychology. Beyond its contributions to academic research, my project stands in support of any efforts in policy and education, to give anger, and emotions in general, an important role. My account of anger's nature and rationality supports the emotion's relevance and importance in the public sphere. Anger should be heard, channeled, and considered, rather than dismissed or repressed. Lastly, my project suggests that some schools of clinical psychology should be more mindful of their patients' social identity, for my project stands in line with the anger of the oppressed having distinctive psychological and epistemic value. My project, therefore, speaks against blanket therapeutic approaches to the anger of agents.

For my parents,
who let me think and feel without constraint.

Contents

Acknowledgements

Before we turn to anger, there is a great amount of gratitude I must express. I am privileged to have been pupil to a number of brilliant educators throughout my life, but three in particular marked my early years significantly. Mrs. Correia, who steered an overactive mind in the best directions, Miss Blanchard, whose tough-love was necessary and support unwavering, and Mr. Smith, on whom I must pause. Mr. Smith was my school's headmaster, as well as my teacher for year 7 English, and year 13 Theory of Knowledge. Under Mr. Smith's excellent instruction these subjects quickly became my favorite, a feat which played a determinant role in bringing me to graduate studies in Philosophy.

Mr. Smith's excellence was equally manifest beyond the classroom, as Headmaster. He put the students first, which I was particularly fortunate of when my continuance at St. Julian's School was threatened. Over the entirety of my secondary school education Mr. Smith indulged my skepticism (often cynicism), took me up in debate on whatever topic was my focus that week, and encouraged me to grow and excel. Without his indulgence I would not be here today.

I am beyond privileged to have been a St. Julian's 'lifer', for which I have to give most thanks to my father. The resources made available there, and the community I grew up with, were indispensable to getting me here.

Dr. Cook designed and ran UCL's unique undergraduate degree in Neuroscience. He designed a historically oriented and, importantly, essay-based degree program, one that privileged argumentative clarity and methodological transparency. Despite Dr. Cook's skepticism of 'philosophy', his program was ideal for a budding philosopher. I thank him for having designed such a superb program, for admitting me onto it, and for supporting me even when it became clear that my interest in neuroscience would remain theoretical.

My conversion to philosophy was undergone via a one-year masters at KCL, where Nick Shea's teaching and encouragement motivated me to continue on to doctoral studies. I thank Nick for supporting my application to graduate school, and for being an exemplar interdisciplinary philosopher mind, and mentor. His support led me back to UCL for the MPhilStud. The transition from KCL to UCL philosophy was not an easy one. The UCL graduate community, at that particular time, was not as diverse or inclusive an environment as I was accustomed to. The culture that predominated there made me doubt my place within it, as well as my aptitude. Things were set to improve, but the actions of a number of people helped me navigate those early times. Maarten Steenhagen was the first to make me feel welcome, as well as to apologize for the gender imbalance of my cohort. Tony Cheng introduced me to the Institute of Philosophy and encouraged me to pursue overlapping interests. At the Institute of Philosophy, I met its deputy-director, Ophelia Deroy who, since examining my MPhilStud thesis, became an indispensable mentor. I thank Ophelia for her continued support, continent-style kindness, and vital career advice.

In the first year of my PhD Amia Srinivasan joined the UCL department. Her time here was short, but the effects ripple on. My first teaching role was on her Feminism and Philosophy course, a topic I knew next to nothing about. I couldn't have asked for a better introduction, as well as guidance, to this literature. I thank Amia for our frequent conversations on the material, where she was encouraging and illuminating, and through which a new and pertinent sense of philosophy revealed itself. Amia's support was pivotal to my confidence as a philosopher, and her course, as well as her own research, had a profound effect on my philosophical interests and practice. This thesis owes a great deal to Amia, as do I.

The person to whom I, and this thesis, are most indebted is Lucy O'Brien. At the beginning of the PhD, Lucy made me take a step back and do exploratory work. This was

scary at the time but, as usual, Lucy knew best. Without this exploratory time, I would not have come to work on the emotions. I'm truly privileged, and grateful, for the time I have spent, at first mostly listening, and later hopefully engaging, such a remarkable mind as Lucy's. One leaves Lucy's generously long supervisions reassured that there is such a thing as Philosophy, and that it is beautiful, and worthwhile. Lucy somehow manages to have the highest standards without letting them defeat the aspirations of her students. Lucy made a philosopher out of me, and for that there is no sufficient thanks.

My time under the supervision of Mike Martin quickly proved my fear of (working with) him unfounded. I am grateful for Mike's time, with which he was always generous, as well as for the care he has for his students and department. I should have realized sooner how much I would gain from working with him. I would also like to thank Ulrike Heuer, and Mark Kalderon, both of whom I worked with during Lucy's sabbatical leave. I am grateful to them for their time, encouragement, and comments on previous drafts of sections of this thesis. I am grateful to Joshua Knobe for supervising me during my time at Yale, and for being one the kindest and most encouraging people I have ever met. Lastly, I would like to thank Christine Tappolet for hosting me in Montreal to work on my final substantive chapter and for making my time there a great experience.

I must of course thank the LAHP, as well as UCL, for generously funding my graduate studies, without which I could not have embarked on this project.

I want to thank my fellow MAP committee members, Catherine Dale, Jessica Fischer and Nikhil Venkatesh for their continued efforts in helping to make our community one that I am proud to be a part of. I hope the Harriet and Helen Memorial Lecture goes from stride to stride, and acts as a reminder of the type of academic environment that should be strived for. Rob Simpson has played an indispensable role in MAPs efforts. I am so grateful to Rob for the time he dedicates these matters, and the care with which he pursues them. I also thank Rob for commenting on drafts of a chapter of the thesis, as well as for his commitment to student well-being, from which I was so fortunate to benefit.

I am very fortunate to have been a part of a vibrant and active research community throughout my doctoral studies. In addition to those I have already mentioned, (I'm sure to have forgotten many) I would like to thank: Léa Salje, Ashley Shaw, Miguel Ferreira dos Santos, Pete Faulconbridge, Jonas Vandieken, Ben Fardell, Julian Bacharach, Ilaria Cozzaglio, Michael Markunas, James Laing, Rowan Mellor, David Olbrich, Alex Geddes, Alec Hinshelwood, Niels Christensen, Naomi Clare, Klara Andersson, Oda Ottosen, Kirstine La Cour and Bex Rowson. Special thanks to Mog Hampson, Showkat Ali, Vanessa Carr, Polly Mitchell and Taylor Enoch, as well as Tom Williams and Tim Short. Thanks to Andrew Knox for his exemplar friendship. Particular thanks to Charles Jansen and Edgar Phillips for being my pals through it all, and for reading and discussing sections of the thesis.

The final department person I want to thank deserves more than I can here express. Richard Edwards, thank you for saving my life more than a few times, and all the many many things you do that keep the department running. I hope you know how much you are appreciated.

I must thank some special people for their friendship. Para além das tugas londrinas, I thank Leonie Walker, Amber Fahey, Vicky Bright, Mariana Siqueira, and Anna Kolesnikova for sticking with me over the years.

Lastly, agradeço à minha família, sem quem nada disto faria sentido. My brother and best friend, Manuel Silva, thank you for your unconditional support, e sobre tudo, por seres quem és. For my parents, Jorge Nuno Silva and Raquel Sousa, words won't do. A eles lhes dedico esta tese, numa aproximação necessariamente falhada da gratidão e do amor que lhes tenho. Por fim, agradeço ao Alexandre, por ser o homem com quem eu jamais ousaria sonhar.

'Como uma criança antes de a ensinarem a ser grande,
Fui verdadeiro e leal ao que vi e ouvi.'

- Fernando Pessoa, 'Fragmentos'

'Like a child before having been taught to grow up,
I was true and loyal to what I saw and heard.'

- Fernando Pessoa, 'Fragments'

Introduction

As the *Iliad* begins with the singing of Achilles' rage (*mênin*), it is often said that 'anger' is the first word in western literature[1]. Anger drives 'revenge drama' from antiquity to contemporary film. In Tarantino's *Kill Bill*, much as in epic poetry, the protagonist's anger is presented as justified, and the work is an exercise in depicting anger run free in pursuit of its revenge. In such drama, anger leaves a trail of bodies and blood.

It is a prevalent and influential view, throughout both the literary and philosophical canon, that anger is, at its core, vengeful and violent[2]. The aim of this thesis is to challenge this widespread traditional view of anger, and to provide an account of anger's rationality. 'Rationality' has been said to be 'one of the slipperiest terms in the philosophical lexicon' (Plantinga 1993: 133)[3]. This might make us doubt the term's theoretical utility[4]. The term will, however, prove useful to us. It will, in the first instance, help us understand much of what has been said about anger in the philosophical literature, as claims pertaining to its (ir)rationality. This will, in turn, help us identify, and make sense of, specific points over which there is disagreement in the literature.

I do not propose one working notion of rationality, but rather, three distinct notions. What these three notions have in common is that they all relate to ways in which anger is often considered to be irrational. Briefly, by *normative rationality*, I will mean anger's ability to respond to reasons for it, by *epistemic rationality*, I will mean anger's role in the provision of evidence or reason for beliefs, and by *instrumental rationality,* I will mean anger's role in securing the practical aims of those in anger. There are two ways in which anger can fail to be normatively rational worth mentioning here. It can be arational, that is, not the sort of state that is had for reasons at all. Or it can be normatively irrational, that is, capable of responding to reasons but not reason-responsive at a particular time. Anger is epistemically irrational if it does not play a positive epistemic role in the generation of justified beliefs. Normative and epistemic rationality are intimately related because, I will argue, anger's ability

---

[1]   See for example Burnyeat (2002), Scott (2010), and Srinivasan (2018).
[2]   Ancient supporters of a retributive construal of anger include Aristotle (2004) and Seneca (2010). Contemporary supporters include Nussbaum (2015, 2016) and Pettigrove (2012).
[3]   Plantinga (1993) distinguishes five distinct ways in which the word is used in philosophy.
[4]   Indeed, Goldman (1986) excludes rationality from the terms of epistemic evaluation which he seeks to analyze, on the grounds that 'this notion is so vague in ordinary usage, and so disparately employed by different philosophers and social scientists, that it has limited usefulness' (27).

to play strong epistemic roles will partly depend on features of its normative rationality. Lastly, anger is instrumentally irrational if it typically undermines the pursuit of the angry agent's practical ends, or alternatively, if it is an ineffective way of pursuing them.

If anger is like an itch or feelings of hunger, it can be related to certain causes, but remains arational in that it is not something for which justification can be given or sought. Anger is normatively irrational, in the second sense, when one experiences anger despite lacking reason for it. If in anger, you become resistant to disconfirming evidence against false beliefs, or your anger causes confirmation bias in favour of beliefs for which you lack sound evidence, your anger is epistemically irrational. When anger causes you to act in ways that you later regret, or in ways which you are actively attempting to avoid, it is instrumentally irrational as it undermines your goals. I do not propose these three types of rationality to be exhaustive of anger's rationality. But an account of anger's rationality would, I think, be incomplete without an account of each of these central types.

Throughout the project, I follow many of those that have written on anger in using the term 'anger' to encompass a set of affectively related phenomena[5]. Wrath, fury, and rage, for example, are anger of particularly high proportion or intensity, while resentment, and indignation, have sometimes been used to refer specifically to anger felt towards morally loaded states of affair[6]. I use generic term of 'anger' to encompass all such phenomena[7], as they are forms anger can take.

Philosophical work on anger can be roughly divided into those who condemn the emotion and those who praise it. Recent attacks on anger, by Nussbaum (2015, 2016) and Pettigrove (2012), follow Seneca in condemning the emotion, while feminist and black thinkers have been amongst anger's greatest proponents. In chapter 1, I outline these two stances on anger and argue that the points over which disagreement arises all relate to at least one sense of anger's (ir)rationality. Put simply, those who condemn anger take the emotion to be, in

---

[5] Nussbaum (2016), for example, uses 'anger' as a genus term, and construes specific types of anger as species of the anger genus. Pettigrove (2012) and Cogley (2014) also use 'anger' to encompass anger-related phenomena.

[6] See for example Miceli and Castelfranchi (2017).

[7] I do so for a number of related reasons. First, what all these phenomena seem to have in common is that they are thought of as forms that anger can take. Second, it is common to do so in the literatures I rely upon (see fn 5). Third, it seems hard to cash out distinctions between these different forms of anger that would justify they be given entirely separate treatment to each other. My account of anger's nature will provide support for this last point.

various senses irrational, while those who praise the emotion take it to be, in various senses, rational. The two stances on anger have four main points of disagreement. These are, firstly, whether anger is a hindrance, or an aid, to rational thinking. Secondly, whether anger is practically beneficial to the angry agent, or, whether it is most often counterproductive to securing their aims. Thirdly, whether anger is promoting or diminishing of the psychological well-being of the angry agent and, lastly, whether anger is by its nature tied to harming others or retribution.

Whether anger is counterproductive to the agent's practical goals relates to what I call anger's *instrumental rationality,* in furthering, or undermining, the agent's practical aims. Whether anger is psychologically beneficial to the emoting agent relates to whether anger is instrumentally rational in a specific sense, namely whether it is persevering of the agent's psychological well-being. Whether anger is an aid, or a hindrance, to 'rational thinking' relates both to what I call anger's *normative rationality*, that is, whether, and how, anger is a state capable of responding to reasons, as well as to what I call anger's *epistemic rationality*, that is, whether anger has a positive or negative effect in leading agents to form beliefs based on adequate evidence. Lastly, anger's nature is related to its rationality, albeit indirectly, for our theory of anger will impact what types of rationality the emotion is capable of, and it will have implications for the form distinct types of rationality might take.

After outlining the terrain of the anger debate, I propose that a foundational approach to it is called for. That is, I propose that making the target phenomenon of anger central, and asking basic questions of its nature and potential rationality, will help move the debate forward. In the second half of chapter 1, I propose four desiderata, and two methodological constraints, which should guide a foundational approach to anger. The four desiderata are for:

1. The provision of an account of anger's nature
2. The provision of an account of anger's instrumental rationality
3. The provision of an account of anger's normative rationality
4. The provision of an account of anger's epistemic rationality

My methodological constraints are that in pursuing these desiderata, we remain informed both by the most relevant work in the brain and behavioural sciences, as well as by features of the social reality in which anger unfolds. From chapter 2 onwards, my chapters pursue

each desideratum in turn. Pursuing these desiderata will, I argue, lend support to those who praise anger.

Although we will see that feminist and black philosophers praise anger, their work has not focused on the provision of theories of anger and, therefore, competing accounts that directly oppose the retributive construal of anger of those who condemn it, are lacking. That being said, by endorsing social constructivist theories of emotion, these thinkers indirectly challenge constitutive links between anger and any specific desire. Over two chapters, I fill what I take to be a gap in the literature in support of anger. Over chapters 2 and 3, I provide a systematic and empirically informed investigation into anger's nature. The account of anger I end up with is more realistic than the anger of revenge dramas, or any account of anger as constitutively tied to retribution. I argue that the nature of human anger is complex and heterogenous. This complexity will give us reason to reject the traditional construal of anger as constitutively tied to retribution, and allow room for the constructive role feminists have afforded anger in the practical realm. My account of anger, while remaining empirically grounded, fits well with the social constructivist insights of feminist philosophers. It also defends anger's tie to a desire for recognition, rather than retribution, which is highlighted in the feminist literature.

In chapter 2 I critique perhaps the most influential theory in the affective sciences, Basic Emotion Theory (BET). This theory has been highly influential in the philosophical literature and plays a role in support of a simplistic understanding of anger as a 'basic' affect program that can be used in support of those who condemn anger. I argue that BET does not survive scrutiny. The sense in which anger emerges as 'basic' is a very thin one: that all humans (and presumably some animals) experience anger related phenomena that share a common affect program *ancestor*. Attention to our evolutionary history will emphasize our adaptation to cooperative living, an evolutionary shift that significantly shaped the function of our emotions. We will see that evolutionary and neuroscientific evidence does not support a distinction between basic forms of anger which are hard-wired on the one hand, and non-basic forms of anger which are dependent on cultural contingencies and higher-level cognitive capacities, on the other. Human anger emerges as a heterogenous yet unified phenomena, with a plurality of evolutionary functions, where the link between anger and specific actions will, unsurprisingly, be weakened.

In chapter 3 I focus on the specific claim, upheld by most of those who condemn anger, that anger involves a constitutive desire for retribution, payback or revenge. Although chapter 2 laid the groundwork for skepticism about anger involving one particular function, my argument in chapter 2 is still compatible with anger involving a constitutive desire for retribution that manifests itself in a wide and diverse set of action types. The traditional commitment to anger involving a desire for retribution forms the basis of modern critiques of the role feminist philosophers have taken anger to play in fights for revolutionary justice. Nussbaum (2016) and Pettigrove (2012) are committed to what I call the *inefficacy claim*. That is, that anger is ineffective at fighting social injustice[8]. This is because, as for them anger aims for payback, anger typically involves acting destructively or punitively, which, in turn, invites retaliation from anger's targets. For this reason, these thinkers take anger to typically, or at least often, be counterproductive to the aims of the oppressed individuals who experience it. This amounts to a claim about anger's instrumental irrationality in the domain in question, namely fighting against social injustice. Using empirical work to supplement my argument, I argue that anger is not typically instrumentally irrational, in personal or, political life. This will lead us to reject the constitutive tie between anger and a desire for retribution. Instead, anger will emerge as contingently, rather than constitutively, related to at least two desires: retribution and recognition. A desire for recognition is a desire for epistemic change in the target of anger, rather than for their suffering, it is a desire that they understand that they have committed a wrong.

By the end of chapter 3 we will have taken considerable steps in meeting the first two desiderata. I will have provided a systematic, empirically informed, account of angers nature, which undermines the construal of anger as constitutively tied to retribution. My argument for anger's efficacy in the realm of social justice sheds light on why the retributive construal

---

[8] I use the terms 'social injustice' and 'oppression' mostly interchangeably throughout the project. They refer to acts or situations that are unfair, rather than merely unfortunate. That is, they capture non-accidental harms and they are normative terms in that they say that things should be, or should have been, otherwise. Where I use these terms in slightly different ways, I give social injustices more of an agential gloss, that is, injustices can be seen to be caused by an agent through their action, or omission of action. I use oppression, on the other hand, at times as more of a structural notion. This is in line with conceptions of injustice (see Miller 2017) and oppression (see Young 1988 and Frye 2000) in the literature. In line with this, an agent can suffer an injustice despite not being oppressed. We might think that oppression involves a system of interrelated injustices that target the same group. In this sense, oppression is something stronger or more pervasive than mere social injustice. That being said, it seems felicitous to say that one can live under *conditions* of social injustice, and that agents can *act* in oppressive ways to others. The definition of each of these terms, as well as the relation between them, is the subject of debate amongst philosophers working directly on these issues. I do not wish to take a substantive stand on the proper use or extension of these terms. My project can be seen as a contribution regarding the relation anger bears to both, or either, of these concerns, depending on the particular view of social justice and oppression one endorses.

of anger may enjoy such prevalence. I provide an outline of the conditions under which anger tends to involve one of these desires over the other. Empirical work suggests that the conditions under which anger manifests itself in actions that seek retribution, as opposed to recognition, are those in which the agent has nothing to lose. These are situations where all other options may have been exhausted, and where one's anger may have been systematically dismissed. Individuals living under conditions of extreme social injustice often have little to lose. This means that those who condemn anger, based on its retributive nature, may be making a mistake regarding anger's nature that serves to perpetuate the very oppression anger arose in response to. That is, the traditional construal of anger as constitutively involving a desire for retribution, stipulates a necessary connection between anger and a desire that, on my account, may depend upon 'nothing to lose' conditions for its manifestation. Traditional accounts of anger obscure the constructive potential of the emotion by taking contingent features of anger, in particularly dire circumstances, to characterize its nature.

Although feminist philosophers advocate that anger plays a strong epistemic role, they do not propose specific accounts of anger as reason-responsive or of anger's role in generating justified beliefs. I do just this in pursuing my third and fourth desiderata. To do so, I combine insights from feminist philosophy and contemporary work in moral psychology and philosophy of emotion. These are fields which, unfortunately, are typically pursued independently to each other.

Chapter 4 focuses on the third desideratum. I argue that emotions enjoy a sui generis normative rationality. I begin by outlining two main ways of making sense of reason responsiveness; rationalist approaches, where reason-responsiveness involves compliance with norms of adequate reasoning, and basic reason-relation approaches, where reason-responsiveness involves only a relation between attitudes and facts that support them. Recent accounts of emotional reason-responsiveness, which I call 'Agential Disposition' accounts, have favoured the former sort of approach. These are accounts that cash out emotional reason-responsiveness in terms of *dispositions* to comply with reasoning requirements. I argue that such approaches face a number of problems, and that a reason-basic approach should be favoured instead. After doing so, I develop an account of this sort, which is largely lacking in the literature. Distinguishing reason-responsiveness from compliance with norms of reasoning allows us to isolate cases where anger is appropriate, that is, where there are reasons for it, from cases where, in being angry, one violates specific norms of rationality.

I begin by supporting my construal of emotions as phenomena capable of reason-responsiveness, by appeal to empirical work on attentional dynamics during emotion episodes. This leaves us in a good position to then move on to characterize the basic reason-relation characteristic of emotions. I argue that this reason-relation is sui generis, as it is irreducible to the reason-relation characteristic of the epistemic, or practical, realm. I argue that there are, therefore, not only epistemic reasons, and practical reasons, but also *affective* reasons. There is a distinctive relation between emotional attitudes and the facts that support them. This relation is characterized by a distinctive feature which I call 'strictness'. Affective reasons are strict in that they stand in support of emotional attitudes even when other affective reasons are at play. In other words, affective reasons are strict in that they do not cancel each other out. The same is not true of epistemic or practical reasons. The sui generis reason-relation characteristic of emotions allows them to play distinctive epistemic roles.

In chapter 5 I defend this epistemic role. For anger to be epistemically rational, it must play a role in leading agents to form justified beliefs. There are two ways anger might do this: by motivating agents to look for reasons for their anger, and alternatively, or in addition to this, by themselves constituting sufficient evidence for the justification of beliefs. Here, contemporary philosophy of emotion can help us provide an account of the positive epistemic role feminist philosophers advocate of anger. Note that anger could still turn out to be epistemically irrational if, despite playing one, or both, of the positive epistemic roles just highlighted, it tended to have epistemically noxious effects that could outweigh the positive ones. Those who condemn anger take the emotion to cloud judgement, bias reasoning, and often lead to false beliefs, despite sometimes playing positive epistemic roles. I argue that we have reason to doubt anger's epistemic irrationality for two reasons. Firstly, we will see that anger's noxious epistemic effects have been exaggerated, but secondly, we will see that anger has a distinctive epistemic role to play, that is, a role that might *require* anger for its positive epistemic effects. This means that the bar will be much higher for those who seek to argue that anger is on average, or all-things-considered, epistemically irrational.

The crucial epistemic role that anger is thought to play, is one advocated by feminist philosophers of anger that arises under conditions of oppression as an 'outlaw emotion'. This is when one's anger goes against one's web of internalized oppressive beliefs and acts as one's sole guide to truths regarding one's oppression. We are currently lacking a systematic

account for how anger plays crucial epistemic roles under conditions of oppression. I provide such an account in chapter 5. This account will hinge on anger playing a justificatory epistemic role. I argue that, although anger often motivates us to search for reasons for it, the emotion also has direct justificatory force of its own in support of evaluative beliefs. The epistemic thesis that emotions play justificatory roles themselves, has come under attack in the recent philosophy of emotion literature. The first half of chapter 5 defends the view that emotions play justificatory roles themselves, against objections raised in the literature.

Although the epistemic role of anger as an outlaw emotion fits well with the emotion playing a justificatory role in support of evaluative beliefs, the received account of the justificatory role of emotions actually risks ruling them out. This is because the justificatory role is a defeasible one, and outlaw emotion cases are one's we have reason to think always involve defeaters. Outlaw emotions, as we saw above, are ones that arise under conditions of oppression and that go against the prevailing ideology. We, therefore, need a story for why one's coherent web of internalized beliefs do not defeat the justification of beliefs based on outlaw emotions. The second half of chapter 5 motivates, and then responds to, this unacknowledged objection to the received justificatory role of emotions. I will argue that only an externalist social epistemology can provide an adequate response to the objection I raise and that, therefore, this is the recommended framework for accounting for anger's epistemic role.

Chapter 4 and 5 allow me to take significant steps in meeting my third and fourth desiderata. These chapters develop accounts of anger's normative and epistemic rationality, respectively, and in so doing provide support for anger playing not only strong, but distinctive, epistemic roles. In chapter 6 I offer concluding remarks regarding the account of anger's rationality that I have put forward.

In taking significant steps towards meeting my four desiderata I will have provided a foundational account of anger as a phenomena that is rational in at least three distinct senses. Anger emerges as complex heterogenous state that is prima facie instrumentally rational in the pursuit of one's aims, as well as a state that, in having a sui generis normative rationality, can play distinctive epistemic roles. My project strengthens the case of those who praise anger, by providing a systematic account of anger as rational. In doing so, my project also

more precisely maps the terrain of the anger debate, such that particular points of substantive disagreement can be made out and targeted by those who wish to condemn the emotion.

1

Two-Faced Anger

*'Anger in the latin west was a sin,
but a sin that could be turned into a virtue,
monopolized by an aristocracy.'*
– Rosenwein, in 'Anger's Past'

Our personal and academic stance towards anger is perhaps best characterized as intensely ambivalent. We caution against anger in our personal lives, citing its futility and destructive consequences, yet we remain suspicious of those in whom injustices do not invoke outrage. Condemned as a personal vice on the one hand, while commended as a justified attitude on the other, anger is often cast as a destructive force while at other times hailed as necessary fuel for positive change. The Stoics and many Buddhist thinkers famously advocated against anger (Śāntideva 1997; Seneca 2010; Bommarito 2017), while philosophers interested in oppression have, on the other hand, highlighted the crucial motivational, political and epistemic value of anger in fighting against the oppressive status quo (Lorde 1981; Frye 1983; hooks 1995; Lugones 2003; Fanon 2008; Yancy 2008).

I will outline these opposing construals of anger in turn. We will see that there are four substantive disagreements regarding anger. These are, firstly, whether anger is a benefit or hinderance to rational thought. Secondly, whether anger is practically beneficial to the angry agent, or whether it is most often counterproductive to securing their aims. Thirdly, whether anger is promoting or diminishing of the psychological well-being of the agent and, lastly, whether anger is by its nature tied to harming others or retribution.

After outlining the terrain of the anger debate, I propose four desiderata as well as two methodological constraints for moving the debate forward. I pinpoint four desiderata which form part of a foundational approach to the debate on anger and focus on questions within the metaphysics, epistemology and basic normativity of the emotion. This thesis will take steps in meeting these desiderata and in so doing provides a, currently lacking, foundational approach to the anger debate.

1.1   Anger as Vice

In the *Psychomachia*, one of the first medieval allegories, Prudentius constructs an epic poem where virtues win over vices in seven paired battle scenes. In it, Patience is attacked by anger - *Ira*, who's pike and javelins slide effortlessly off Patience's armor. *Ira* explodes in frustration at her failed efforts. Her fury leads her quickly to a frenzied suicide (Freedman 1998). This 5th century illustration of anger has much in common with a present-day view of anger as vicious: as ineffective at enacting change, as uncontrolled, and ultimately, as counterproductive and self-harming.

In Christian iconography, anger is often depicted in the form of the devil and presented as disheveled and blinded by evil desire. This isn't surprising given that anger features as one of Christianity's seven deadly sins. Indeed, John Cassian, who popularized the seven deadly sins in the west, wrote that 'as long as anger remains.. we cannot be partakers of life, nor can we become partakers of wisdom, nor can we attain immortal life' (Freedman 1998: 13). There is, therefore, an inheritance, at least in the Christian west, of anger being seen as an enemy to reason and virtue.

The writings of perhaps the most famous of anger's critics, Seneca, had a pervasive influence on the fathers of the Christian church (Freedman, 1998)[9]. Seneca's influence is not limited to the Christian west, however. Indeed, the religion most faithful to Stoic teachings is arguably Buddhism. The key Buddhist text, *The Way of the Bodhisattva*, dedicates an entire chapter to anger, arguing that it 'has no other function than that of causing harm' (6.8) (Śāntideva 1997).

For Seneca (2010), anger 'consists entirely in aroused assault. Raging with an inhuman desire to inflict pain' (1.1). Anger is therefore uncontrollable and retributive. Anger is 'hungry for a vengeance that will bring down the avenger too' (1.1), making the emotion often counterproductive to the emoting agent. Indeed, 'angry people curse their children with death, themselves with poverty, their households with ruin, and they deny they're angry just as madmen deny they're insane' (2.36.5).

In likening anger to insanity, anger is cast at complete odds to reason 'distinct symptoms that mark madmen…the signs of angry men too, are the same' (1.1.3). As a stoic, Seneca viewed all passions as essentially defective beliefs that are to be curtailed so as to live a virtuous life. Seneca's *De Ira* casts anger as the worst amongst the emotions, 'see the foundations of the most celebrated cities hardly now discerned; they were ruined by anger' (1.2.1), and councils on how to restrain actions done out of anger. Crucially, however, he provides advice on how to avoid anger altogether, through education and cultivation of other moral virtues, 'we shouldn't control anger, but destroy it entirely—for what "control" is there for a thing that's fundamentally wicked?' (3.42). For Seneca, 'no pestilence has been more costly for the human race' (1.2) than anger.

Seneca gives us three related reasons to condemn anger. Firstly, anger is responsible for violent and destructive actions against others. Secondly, in causing destructive actions, anger is often counterproductive to the aims of the angry agent themselves. Actions done out of anger can compromise one's long term aims, as well as provoke opponents into retaliation against oneself. Lastly, by desiring the suffering of another, anger is plausibly 'fundamentally

---

[9] This is most evident in the writings of Martin de Braga, who popularized a self-authored book entitled *De Ira* made up almost entirely of verbatim quotations from Seneca's work by the same title. This early medieval work resurrected the strong anti-anger sentiment of the original text and spread its message throughout the growing Christian Kingdom. By the late middle ages, monastery rules prohibited anger of monks, and religious advisors to European monarchs counselled against anger, highlighting its incompatibility with a virtuous leader (see Freedman 1998)

wicked', whether or not it causes destructive actions. Seneca's *De Ira* is a dialogue that records a debate between a Stoic and an Aristotelian view on anger. The Aristotelian view allows that virtue sometimes call for anger, and Seneca is engaged in refuting this view (see Cherry and Flanagan 2017). Seneca says that his Aristotelian opponent can raise the following objection to his total condemnation of anger: 'Just as virtue is kindly disposed to honorable behavior, so should it greet disgraceful behavior with anger' (2.6), to which Seneca replies that 'virtue will never make the mistake of imitating vices… It's a natural property of virtue to be glad and joyful' (2.6.2). Seneca rules out the possibility of anger ever being called for, or being justified, even if it's negative consequences might be able to be controlled. Anger, for Seneca, is just at odds with a virtuous life.

Aristotle's view, on the other hand, goes some way in capturing the two-faced features of anger, allowing anger to enter the realm of virtue. Aristotle writes that:

> There is praise for someone who gets angry at the right things and with the right people, as well as in the right way, at the right time and for the right length of time. This, then, will be the even-tempered person, since it is his even temper that is praised (*NE* 1125b32-33).[10]

This has been read as an endorsement that anger is sometimes virtuous, when it is apt and in proportion. Failing to become angry when a situation merits anger emerges as a potential moral failure in its own right; 'people who do not get angry at things that they ought to get angry at are thought to be foolish' (*NE* 1126a4). In stark opposition to Seneca, anger is not only permitted, but at times required. The appropriately angry individual is worthy of praise. On such a reading of Aristotle, there is virtuous anger.

Reading Aristotle to cast such a positive light on anger is, however, not without its difficulties. For one, he writes that it is the appropriately angry person's 'even temper' that is praised, rather than anger itself. This suggests that it is not anger that is virtuous but the agent's ability to control it. Furthermore, Aristotle's definition of virtuous anger raises a number of problems for the idea of virtuous anger. Aristotle defines anger as follows:

> Anger (orgê) may be defined as a desire accompanied by pain for a conspicuous revenge, on account of a perceived slight on the part of people who are not fit to slight one or one's own (*Rhet.* 1378a31-33).

---

[10] I use the abbreviations *NE* and *Rhet.* for the *Nicomachean Ethics* and the *Rhetoric*, respectively.

This highly influential definition comes into two potential problems, firstly, it ties anger conceptually to a desire for revenge. Secondly, the last part of the definition suggests that social status is an important determinant of whether anger is justified (Cherry and Flanagan 2017). I discuss the latter before turning to the former. Aristotle writes:

> nor is it possible for him to be angry with someone more powerful than himself, since here the appropriate emotion is fear, and it is impossible, after all, to be angry and afraid at the same time (*Rhet.* 1380a).

These observations have lent support to a popular reading of Aristotle as providing an Honour or Status account of virtuous anger (see Stocker and Hegeman 1996, Nussbaum 2016). Konstan (2007) writes that 'For Aristotle, what counts as belittlement [i.e. slight] depends on status: if your position is inferior, it is no shame to be reminded of it' (55). This suggests that anger is not apt when directed at those with higher social standing than oneself, anger is justified only when directed at 'people who are not fit to slight one'. This type of Status reading makes ones standing within a social hierarchy central to whether anger is apt. On such accounts, anger has been said to be tied to 'narcissism' which might itself be condemnable (Stocker and Hegeman 1996: 269). In any case there seems to be an asymmetry to anger's justifiability; when anger is directed at those of lower social standing it can be apt, while anger flowing in the opposite direction is ruled out as unjustified, and perhaps even incomprehensible.

This asymmetry is mirrored in the history of the Latin west. In Christian scripture God can become enraged, and justifiably so. As supremely virtuous, God's anger does not threaten his divinity and instead anger is made acceptable and expected amongst the almighty. This had its counterpart in the earthly realm, where *ira regis*, the anger of Kings, was accepted as rational and virtuous, likely through a divinely inherited legitimacy, that the Christian church granted the monarch. Kings, therefore, through their proximity to God, and their divine right, were also entitled to anger. Virtuous anger soon became the purview of the noble alone, extending beyond the King but excluding the lower classes, 'anger in the latin west was a sin, but a sin that could be turned into a virtue, monopolized by an aristocracy' (Rosenwein 1998: 5).

Indeed, the anger of peasants has been historically either ridiculed for being mistaken, or vilified for being violent and animal like; 'comic or murderous, peasant anger was quintessentially irrational' (Freedman 1998: 171). As anger was more than often about

defending one's honour, the anger of peasants seemed inherently misplaced, as they were not allowed any honour that could be in need of defense. Indeed, a recurring satirical depiction of peasants in the middle ages was that of a foolish peasant attempting to pass for a knight, but giving himself away in his failure to act chivalrous or defend his honour in battle (Freedman 1998).

We, therefore, have reason to think that notions 'virtuous anger' in part stem from, and have perpetuated, oppressive social relations. 'The distinction between good and bad anger.. had wide ranging implications for the nature of social and hierarchical bonds' (Barton 1998: 160). This will be important when considering the liberatory and transgressive roles attributed to anger by those who give the emotion a central role in radical politics. It is clear that within the picture just sketched, the anger of a member of a subordinate group could be a revolutionary act, one of demanding honour that the social structures in place presently deny them.

Perhaps the Status or Honour reading of Aristotle's account of anger gets things wrong in situating status within social hierarchies, however[11]. Perhaps what is actually at stake is one's moral status (see Christensen 2016). One is, on such an account, appropriately angry when one's moral worth has not been respected, where one's moral worth is a fact independent of social standing. Such a reading would avoid the asymmetry that previous Status accounts are committed to by allowing subordinate groups to be justifiably angry at those in power. Even if Aristotle's view can be read to avoid a commitment to this problematic asymmetry, however, the view risks ruling virtuous anger out not on account of status based 'correctness' conditions but on account of anger's very nature.

For Aristotle, anger's nature is characterized partly by a desire for revenge. This desire is a desire to inflict pain (*Rhet.* 1382a8, 1382a14-15). Enacting retribution over one's offender satisfies one's desire for revenge as anger ceases 'when he pays back for the offense; for revenge stops anger, as it produces pleasure in return for pain' (*NE* 1126a21-22). Nussbaum (2015; 2016), a modern defender of Aristotle's conception of anger, believes that oppressed groups *can* feel appropriate anger at their oppressors. Nonetheless, Nussbaum has argued that as anger's nature involves a conceptual tie to payback or retribution, the emotion is either inherently immoral or wholly irrational. In chapter 3 the details of Nussbaum's view

---

[11] See Christensen (2016) for a full critique of the Honour or Status reading of Aristotle on anger.

will be considered, for now I wish only to highlight that even Aristotle's view, which many have read as permissive of virtuous anger, faces difficulties on two accounts: firstly, by postulating asymmetrical correctness conditions that would rule out the appropriate anger of subordinate groups, and secondly, by endorsing a retributive conception of anger's nature which risks making the emotion inherently, or at least typically, morally problematic.

Aristotle's commitment to anger sometimes being justified amounts to a claim that it is virtuous at times to desire revenge. Indeed, Seneca writes that 'Aristotle's definition is not very different from ours: he says that anger is the strong desire to return pain for pain' (1.3.3). But this desire is part of what makes anger problematic for Seneca, as well as Nussbaum (2016). The retributive construal of anger, which Nussbaum endorses and calls the 'traditional' view of anger, is alive and well in modern scholarship. Ben- Ze'ev (2000: 384) for example writes that 'the urge to attack is essential to anger, even if it is expressed in a nonstandard aggressive act'. Similarly, Pettigrove (2012: 358) takes anger to 'involve the desire to lash out at its object or to see that object hurt', while for Nussbaum (2015: 41) 'the idea of payback or retribution—in some form, however subtle—is a conceptual part of anger'.

So far, we have seen a strong anti-anger vein hailed by the Seneca, and a number of major religions, as well as a less damning view on the emotion advocated by Aristotle. Aristotle's view seems to allow anger some role within the rational mind and virtuous life, and in so doing is at odds with the Stoic view. The Aristotelian view has been deeply influential in the folk psychology of the Christian west, as well on anglophone academic philosophy. We tend to think that some things merit anger, and that others do not. Our Aristotelian inheritance has not led to educational and societal trends that favour the cultivation of rational anger, however. We still advocate control of anger over the cultivation of virtuous anger (see Stearns and Stearns 1986). Anger's role in rational and virtuous life is still strictly limited. Aristotle's conception of anger is one of anger itself being conceptually tied to vengeance or payback. As long as such a view of the nature of anger is maintained, the emotion's compatibility with an ethical life will remain tightly limited, as seeking to inflict pain on others, or even merely desiring to so, risks being morally problematic.

Another reason to advocate against anger is that it is taken to be an inherently negative experience. Anger is classified as a negative emotion in that it is an unpleasant experience for the angered subject themselves. Anger is said to have negative valence, due to involving a

negative hedonic tone or subjective feeling (Kolnai 1998; Frijda 2005). Recall that for Aristotle, anger is a 'pain' accompanied by a desire for revenge. Anger isn't typically taken to be a pleasant or desirable state to be in, indeed the discomfort caused by anger may be part of its motivational force, pushing the agent to change the situation so as to be able to move out of a painful state. As an inherently negative state, anger can have negative psychological effects, such as being mentally depleting or exhausting, as well as inducing other negative states such as hatred or guilt. The view that frequently being in anger can be corrosive to ones well-being, and negatively affect your mental, and even physical health, is widespread (see Silvermint 2017). Distinct from anger's consequences on other's then, the emotion can be seen as having negative consequences on the emoting agent as well. Both by proving practically counterproductive to securing their aims, and by being psychologically burdensome. Being in anger puts the agent in a negative subjective state, that can cause further negative psychological effects.

Those who condemn anger tend to employ arguments of consequentialist and/or comparative kind, using historical and/or experimental evidence as support. Consequentialist arguments, as we have seen, focus on highlighting the negative and destructive effects of anger and take these to be reasons for minimizing the occurrence of the emotion and for its status as a vice.

Comparative arguments still focus on anger's consequences but compare and contrast them to that of other emotions. Religious writings for example highlight the comparative disadvantage of anger as opposed to virtues such as patience and love. A contemporary advocate of such a comparative approach can be found in Pettigrove (2012), who argues that meekness is the preferable response to offence, even in cases of systemic oppression. Pettigrove cites evidence in support of anger's comparative flaws, including that anger lower's subjects risk evaluations, making those angry more likely to engage in risky behaviour (Lerner and Keltner 2001), and that angered subjects are more likely to take punitive actions against those they blame (Lerner et al. 1998). Much like Pettigrove, Nussbaum (2015; 2016) specifically argues against the constructive roles feminists would have anger play in revolutionary politics. She argues from historical examples that successful social change follows from calm, rather than enraged approaches, and that positive emotions of love and compassion should, therefore, replace anger.

Anger's negative consequences are central to arguments condemning anger. We have seen claims that anger not only causes destructive actions directed at others, but is also psychologically noxious to the angry agent themselves. A further consequentialist line of argument targets not the psychological or practical effects of anger but it's *epistemic* effects. This is present in the ancient and medieval portrayals of anger as akin to madness, and incompatible with reason. Anger here is seen as epistemically noxious and undermining of sound reasoning. Experimental work in psychology is now often cited in support of such a view of anger. The work cited includes observations that angered subjects are more likely to attribute negative events to the intentional actions of other agents, rather than to causes that are not agent related (Keltner et al. 1993), and that angered subjects are more prone to employing stereotypes (Bodenhausen et al. 1994). Angry individuals have also been found to overestimate the likelihood of positive personal outcomes, and underestimate the likelihood of negatives ones, compared to non-emotional subjects (Lerner and Keltner 2001). Angry subjects are also reportedly quicker to attribute negative, than positive, traits to members of an out-group (Lerner and Tiedens 2006). Finally, angered subjects are less likely to trust others (Dunn and Schweitzer 2005)[12].

There are, therefore, five main reasons for advocating against anger: that it has destructive practical effects that harm others, that it often ends up harming the angry agent himself, that anger by its nature involves a desire to harm which should itself be avoided, that anger is painful for those experiencing it, and that it is epistemically detrimental, and is a hinderance on sound reasoning.

We will see that there are those who have a much more positive construal of anger. They do not deny that anger often causes destructive actions, but disagree with those who condemn anger on all four other accounts.

1.2   Moral Anger

Anger has been considered *the* essential political emotion, construed as the paradigmatic response to injustice (Spelman 1989; Holmes 2004; Thompson 2006). It is, therefore, unsurprising that black and feminist philosophers have been amongst anger's strongest

---

[12] These studies have been cited in either Pettigrove (2012), or Cogley (2014), or both.

defenders. As an appropriate response to injustice, 'anger is loaded with information and energy' (Lorde 1984: 127), and hailed as having epistemic, moral and motivational value indispensable to the fight against social injustice. We will see that such thinkers take anger to be useful in both clarifying the nature of this oppression, and in fighting against it. Anger's value as a political emotion involves disagreement with those who condemn anger on four main points: that anger is not of epistemic value, that it is psychologically undesirable, that it is counterproductive to securing the aims of those in anger, and that anger is, by its nature, necessarily tied to retribution. I deal with these in succession.

1.2.1 Epistemic Value

The view that emotions have a strong epistemic role is widespread in the feminist literature (Frye 1983; Friedman 1986; Narayan 1988; Jaggar 1989; Bell 2009). Anger in particular is seen as a knowledge conducive state capable of responding to adequate reasons. Indeed, emotions such as anger are often highlighted in the literature as the only means to gain such epistemic insight:

> A woman who has been taught that a 'woman's place is in the home' may be driven to question this maxim precisely in light of her persistent dissatisfactions and repeated urges to flee from the responsibilities and limitations which structure her domestic life… if her highest principles themselves also include notions of 'appropriate' sex roles, duties to others and the importance of self-sacrifice as an ideal of femininity, then there is not much available among her highest principles to afford an independent standpoint for assessing the maxim about woman's place. Her frustration, grief, and depression, and the motivations to change her life which spring from these sources, may be her only reliable guides. (Friedman, 1986: 31)

In a similar vein, Jagger (1989: 166) writes that 'the hegemony that our society exercises over people's emotional constitution is not total'. The thought is that emotions are capable of tracking reasons that would otherwise be unavailable to us. Jagger coined the term 'outlaw emotion' to refer to these emotions that occur under conditions of oppression, which are 'distinguished by their incompatibility with the dominant perceptions and values' (1989: 166). The capacity to track such reasons is thought to be affected by social standing and oppressive relations;

> the social situation of such people (the oppressed) makes them unable to experience the conventionally prescribed emotions: for instance, people of color are more likely

to experience anger than amusement when a racist joke is recounted, and women subjected to male sexual banter are less likely to be flattered than uncomfortable or even afraid (Jaggar 1989: 166).

We need not be committed to the implausible view that emotions *always* track reasons, or that they always do so better than our belief systems, so as to accept the epistemic value of anger. We need only hold that sometimes emotions better track reasons than our belief systems do. To deny this weaker claim would be to deny the genealogy of liberatory movements, and the role of consciousness raising, where collective efforts to make sense of emotional experiences shared across groups, led to critical progress (see Fricker 1991, Thompson 2006).

In addition to directly tracking wrongs, anger can also be epistemically beneficial in an indirect sense[13]. By observing when and where one's anger is accepted as a legitimate claim about how things stand in the world, one can gain epistemic insight about one's oppression. One's anger will be taken seriously only when it is experienced over concerns that one is perceived as legitimately having a stake in. For example, a women's anger might be taken seriously in the kitchen, an area that is her domain, while being entirely dismissed in the public or political sphere. By observing how one's anger is received by others, Frye (1983) says that 'anger can be an instrument of cartography' (94), in allowing us to map the limits of our agency and worth under conditions of oppression.

Here we have our first point of disagreement with those who condemn anger. As opposed to being epistemically detrimental, feminist philosophers hold anger to be a crucial, and often indispensable, source of knowledge regarding injustice. Those who condemn anger would have anger be a hinderance to clarity and reasoning rather than a guide to important truths. At the very least, they would be committed to taking anger's overall epistemic effects to be negative, even if they granted that anger sometimes cues agents to truths. Pettigrove for example writes that 'even if "moral" anger has the epistemic merits that advocates have claimed on its behalf, these are accompanied by enough epistemic liabilities to temper whatever enthusiasm we might have felt for it' (2012: 364).

---

[13] This point is also made in Bell (2009: 168).

1.2.2 Psychological Value

Oppressed individuals are taken to be in a double bind when it comes to the psychological effects of anger. On the one hand, feminist and black thinkers clearly take anger to be the appropriate response to injustice, on the other, living a life filled with rage is unlikely to be pleasant or desirable. Defenders of anger have been careful to highlight the ambiguity of their position, Audre Lorde, for example writes, 'My anger has meant pain to me but it has also meant survival' (2007: 132).

bell hooks (1995) further highlights anger as a positive psychological act:

> Racial hatred is real. And it is humanizing to be able to resist it with militant rage…When we embrace victimization, we surrender our rage. My rage intensifies because I am not a victim. It burns in my psyche with an intensity that creates clarity. It is a constructive healing rage. (17-18).

Here we have a portrayal of anger as associated with psychological health rather than illness. Anger is also associated with respect, for through her rage hooks resists the internalization of an oppressive self-image. Victims of oppression have highlighted the self-affirming effects of anger. Frederick Douglass (2009), for example, describes the moment he followed his rage and confronted his master who was in the act of 'disciplining' him:

> It rekindled the few expiring embers of freedom, and revived within me a sense of my own manhood. It recalled the departed self-confidence, and inspired me again with a determination to be free. (78)

In a similar vein, critical race theorist George Yancy (2008) writes of confronting the white gaze:

> I am angered. Indeed, I find her gaze disconcerting and despicable. As I undergo this double consciousness, my agency remains intact. My sense of who I am and how I am capable of being...has not been eradicated. I know that I am not a criminal or a rapist. At no point do I either desire to be white or begin to hate my dark skin. (847)

We see here the self-affirming effect of anger, which allows victims of racism to reject the internalization of an oppressive self-image[14]. In *Black Skin, White Masks*, Fanon (2008) writes of a similar encounter:

---

[14] See Leboeuf (2017) for a more detailed phenomenological account of anger in the psychology of the oppressed.

> I can feel the familiar rush of blood surge up from the numerous dispersions of my being. I am about to lose my temper. The fire had died a long time ago, and once again the Negro is trembling. Look how handsome that Negro is! The handsome Negro says, 'Fuck you, madame.' (94)

Anger allows the oppressed to resist and reject the internalization of noxious conceptions of themselves. In becoming angry, one takes oneself to be worthy of respect, and crucially to be the offender's equal. This is a fact which, in situations of oppression, is typically denied. Spelman (1989) writes that:

> If he is in other ways regarded as my superior, when I get angry at him I at least on that occasion am regarding him as no more and no less than my equal. So my anger is in such a case an act of insubordination: I am acting as if I have as much right to judge him as he assumes he has to judge me. (266)

There is therefore psychological value in experiencing anger, for it promotes or involves a positive conception of oneself. Not only can there be pleasure in rejecting an oppressive conception of oneself, and affirming oneself as worthy of respect, but this is also arguably much healthier than not doing so. Internalizing oppressive conceptions of oneself can be not only painful but damaging to one's life prospects. We, therefore, have first-hand accounts of the therapeutic benefit of anger in enabling psychological resistance against oppressive ideology.

In being psychologically beneficial, in the manners just outlined, anger might itself be a state that is not entirely negative to feel. If the very feeling of anger involves asserting one's moral worth, then anger might not be as painful an experience as those who condemn it would have it. Indeed, there is psychological evidence that anger is actually an emotion with mixed, positive and negative, valence. Appraisal theorists take anger to involve a negative appraisal of a triggering event, as well as a positive appraisal regarding one's ability to cope with, or change, the triggering event (Scherer 2005; Lerner and Tiedens 2006). This suggests that even in situations other than oppression, anger is not an entirely negative experience. From the first-personally accounts of individuals living under oppression, we have reason to think that, in appraising the situation as offensive or wrong, anger plausibly also involves a positive appraisal of oneself. We, therefore, have reason to doubt that anger is typically an entirely negative subjective experience.

Besides the effects mentioned, sharing one's anger with a group of individuals can have clear benefits in promoting feelings of group membership and in increasing one's self-confidence (Jasper 2014). There are, then, multiple psychological benefits to experiencing anger, as well

as multiple reasons to think the emotion can be a partly positive experience. Those who condemn anger have said little about these beneficial psychological effects. This might be because they have failed to focus on the personal experiences of members of oppressed groups. The arguments employed by those who condemn anger, however, leave open the possibility of acknowledging these beneficial psychological effects yet remaining committed to an all-things-considered judgement that anger's negative psychological effects outweigh its positive ones. Alternatively, they can argue that other emotions such as love or meekness would be preferable responses to anger triggering circumstances (see Pettigrove 2012; Nussbaum 2016). A story will need to be told for how other emotions can deliver the psychological benefits anger seems to promote. Without the self-affirmation anger seems to involve, critical consciousness and opposition is unlikely to develop. Hence, the importance of anger in the personal psychology of the oppressed is intimately linked to its political potential.

1.2.3 Practical Value

Political activists around the world cite anger as a crucial motivator for political action[15]. In line with this, one of the most crucial roles afforded anger by feminist and black thinkers is its motivational role. Lorde (1997) writes that:

> Every woman has a well-stocked arsenal of anger potentially useful against those oppressions, personal and institutional, which brought that anger into being. Focused with precision it can become a powerful source of energy serving progress and change. And when I speak of change, I do not mean a simple switch of positions or a temporary lessening of tensions, nor the ability to smile or feel good. I am speaking of a basic and radical alteration in those assumptions underlining our lives. (280)

Here anger is characterized as powerful fuel for critical change. Anger spurs us to act and is ubiquitously acknowledged as a potent motivational force. Anger's critics do not deny the emotion's clear motivational power, what they deny is that actions done in or out of anger will actually lead to substantial change. Actions done out of anger, on their account, are uncontrolled and destructive, and risk being ineffective or even counter-productive to the

---

[15] Hessel's (2012) '*Indignez-Vous!*' characterizes indifference as the worst response to injustice, and encourages readers to become outraged at the social injustices of our time: finance capitalism in particular. Hessel's pamphlet inspired the Spanish anti-austerity movement *Los Indignados* whose actions led to a historic turn in national politics, as well as having catalyzed world-wide political activism, including Occupy Wallstreet (see Rosenmann 2012).

angry agent. The thought is that acting destructively is likely to leave the angry agent worse off, either by causing material damage that harms oneself, and one's group, or by triggering retaliation from those at whom anger is directed.

The danger of angry actions is not minimized by feminist writers. After all, Lorde says anger must be 'focused with precision', she also writes that 'My response to racism is anger. I have lived with that anger … learning to use it before it laid my visions to waste, for most of my life' (2007: 124). It is clear that anger's potential perils are not being dismissed, nor is anger being hailed as necessarily righteous. hooks (1995) writes that:

> we learned when we were very little that black people could die from feeling rage and expressing it to the wrong white folks. We learned to choke down our rage ... Rage was reserved for life at home—for one another. (13-14)

The risk of anger's counterproductive effects is, therefore, at the forefront of these thinkers' minds. Nevertheless, they advocate against its suppression:

> By demanding that black people repress and annihilate our rage to assimilate… white folks urge us to remain complicit with their efforts to colonize, oppress, and exploit (hooks 1995: 16).

Despite acknowledging the negative effects anger can have in motivating destructive actions, there thinkers seem to take anger to have constructive potential in guiding meaningful action as well. By not being aimed at a reversal of positions, and instead, aiming for 'progress' and 'change', anger can motivate constructive, as opposed to destructive, actions. These will include actions such as protesting and conscious raising, for example. This is the first sense in which anger can be of practical and political value that I wish to highlight. That is, where anger's political efficacy does not seem to be tied to destroying or attacking the other, in any straightforwardly aggressive sense.

This departs from the destructive construal of anger advocated by those who condemn the emotion. That being said, commitment to the traditional construal of anger as triggering uncontrollable destructive actions is still compatible with it having political value. After all, violent revolutions have been one of the most effective ways of overthrowing the prevailing political order. Assuming anger is a key motivator for such revolutions, destructive anger is of clear political value, in so far as we are concerned with efficacy, rather than morality.

Violent revolutions often have, at least short-term, downsides to those who fight them, however. So perhaps anger is counterproductive in the short-term to those agents that act violently out of it, even if such action leads to important social change. There would, in such cases, be a contrast between the political value of anger, in the long-term, as opposed to the short-term. Alternatively, the contrast could be between who benefits from the political value of anger, rather than when these benefits are felt. That is, anger could prove counterproductive for the angry agent, who, say, sacrifices themselves for a cause, but be of immediate political value to others.

In line with this, it has been argued that it is virtuous to become angered in situations of oppression, despite anger being counterproductive to angry agents. The thought is that anger will promote social change that others will be able to benefit from. Tessman (2005), for example, seems committed to such a view. Tessman (2005) makes the proposal that anger is amongst the 'burdened virtues'. These are dispositions that deserve the title of virtue despite undermining the flourishing of the agents who manifest them. Anger, on this account, is incompatible with personal flourishing given the negative effects it has on angry agents. Being dangerous to manifest in liberatory struggles due to typically provoking retaliation, for example. Anger can result in incarceration and even death. It can, then, be counterproductive, to angry agents living under conditions of oppression to act on their anger. Yet, anger at injustice is virtuous, for Tessman (2005), because this anger can lead to social changes that will promote the flourishing of others[16].

This sort of picture makes some sense, liberatory struggles do, after all, often take place over generations, and people sacrifice themselves for such causes. Anger's practical value may, then, lie in its long-term effects on agents other than the angry agents themselves. Anger's practical value is not limited to its long-term effects, however. Anger can sometimes motivate aggressive revolts that cause an immediate, or at the very least quick, shift in power. Take the famous example of the Haitian 1791 slave revolt; 100,000 slaves revolted against their masters, and raped, tortured, pillaged, mutilated and killed most of the white population (around 4000 people). By 1792 they had control of one third of the island, and by 1794 slavery had been formally abolished. Clearly, in so far as revolutions are spurred by anger,

---

[16] Tessman (2005) has been critiqued by Bell (2009), who argues that anger's virtue should not be understood in consequentialist terms at all. All that matters for anger to be virtuous, for Bell (2009), is that it get things right. Appropriate anger is therefore virtuous anger on her account. In what follows I do not adopt a virtue ethics framework, but I agree that anger's aptness should be distinguished from his effects. Srinivasan (2018), similarly takes a non-instrumental approach to evaluations of anger's value. Being appropriately angry is intrinsically valuable on such views.

the emotion can be effective in overturning social orders on a much shorter time-scale than that of generations. Many of those who revolted clearly benefited from the political value of their actions.

Endorsing a traditional construal of anger where the emotion motivates destructive actions is, therefore, compatible with its political value. There are three overlapping options for the political efficacy of destructive anger: it can lead to a short-term worsening but long-term bettering of the political situation for all relevant agents, including those that acted on their anger, it can lead to social improvements that agents other than the angry agent(s) get to benefit from, and finally, anger can lead to benefits that the angry agent(s), as well as others, benefit from even in the short-term. The efficacy of violent retributive anger should not be overlooked. History has shown stark societal changes to often be prompted by violent means, often fueled by angry masses[17]. There is room to make the case for the justifiability, and, certainly, the efficacy, of even intensely destructive retributive anger. Making such a case will not be my concern, however. Instead, it will be with providing a defense of anger as politically effective, that does not depend on anger being tied to destructive action tendencies at all.

We saw that the claim that anger motivates destructive actions is one endorsed by those who condemn anger and tie the emotion to a constitutive desire for retribution. Those who praise anger have not denied this, although in their defenses of anger's political value they do not focus on its destructive effects. Anger's connection to punitive actions is an empirical claim that has not been systematically challenged in the feminist literature. My concern will be with providing a critique of this empirical claim, and highlighting its theoretical importance. I do so in chapter 3. Such a critique will support the view, present in the feminist literature, that anger's political value is not limited to its destructive potential. In doing so I will provide a currently lacking head on challenge to the empirical assumptions of those who condemn anger.

---

[17] More on this in chapter 3.

1.2.4 Anger's Nature

So far, we have seen that philosophers concerned with oppression have taken anger to have crucial motivational, epistemic and psychological benefits. What about anger's nature? Have black and feminist writers endorsed theories of the emotion that depart from the traditional construal of anger as inherently tied to payback? Yes and no. Providing necessary and sufficient conditions for the nature of emotions has not been among the primary concerns of feminist philosophers. Lutz (2002) writes that:

> While the approaches taken to emotion by feminists have been varied, they share critical and pragmatic purposes. They are critical in the sense that they redefine what is worth knowing about emotions, asking new questions and questioning the interests served by the old questions. They are pragmatic in the sense of aiming to apply the new questions and their answers to benefit in women's (and men's) lives. (106)

Indeed, feminist philosophers have been much more preoccupied with the political dimension of anger than its nature. As we saw with Frye (1983), for example, the focus is on providing 'cartographies' of oppression by observing when and where women's anger receives 'uptake' i.e. is received as a legitimate claim regarding a state of the world. Whenever anger is denied uptake we are witnessing a domain in which female agency is denied. Similarly, hooks focuses on how 'white people remain unable to hear black rage' (1995: 12), and hence unable to recognize racial oppression. Spelman (1989: 270) writes that it is in the interest of those dominant not to make anger available to the oppressed, 'the systematic denial of anger can be seen as a mechanism of subordination, and the existence and expression of anger as an act of insubordination'.

Despite this lack of focus on providing theories of the nature of emotions, feminist philosophers have leaned towards social constructionist theories (see Jagger 1989, Fricker 1991). Jaggar (1989) for example writes:

> The apparently individual and involuntary character of our emotional experience is often taken as evidence that emotions are presocial, instinctive responses, deter-mined by our biological constitution. This inference, however, is quite mistaken… mature human emotions can be seen neither as instinctive nor as biologically determined. Instead, they are socially constructed on several levels. (157)

Emotions are thought to be shaped by upbringing and culture. Such commitments oppose universal construal's of anger's nature, which in itself stands in opposition to the traditional view of anger by denying that the emotion is necessarily tied to any desire, be it retribution

or otherwise. The literature has, in this way, questioned the traditional construal of anger as involving a close link to retribution, indirectly, by reframing which questions are worth asking about the emotion, as well as by refusing to attempt a formulaic definition of anger independent of cultural context.

In line with Aristotle's conception of anger however, feminist philosophers agree that anger is minimally cognitive. Feminist philosophers, much like Aristotle, take anger to be an intentional phenomenon that is responsive to wrongs and subjectable to appropriateness conditions (Jaggar 1989; Fricker 1991). Indeed, this is implicit in the political, epistemic and psychological roles they afford anger. As opposed to Aristotle however, who takes the objects of anger to be limited to particular individuals, and who holds that the emotion cannot be directed at 'mankind' as a whole (*Rhet.* 1378a), feminist philosophers take anger to admit social objects, such institutions, and social states of affair, such as 'sexism' (Frye 1983; Jaggar 1989; Fricker 1991).

The constitutive link between anger and desire for revenge has been directly questioned in recent writing. An alternative conative component of anger has been highlighted in the feminist literature: a desire for recognition. Srinivasan (2018) for example writes that anger often aims 'to get the perpetrator to recognise the pain she has caused me, the wrong she has done me', rather than to make the target of anger suffer. Similarly, Cogley (2014) writes that, in anger, an agent:

> confronts the target of her anger in an attempt to bring the target's attention to her cause for anger. She then asks after or demands an explanation or justification; if the justification is insufficient, she acts to change the situation. (211)

On such a picture, anger aims, in the first instance, at an epistemic change in its targets, that they come to appreciate or recognize the wrong they have caused. And it might be only when this aim is unmet, or blocked, that anger then aims to punish its target.

The view that anger is more concerned with issuing demands for respect than punishing its targets, is common amongst moral philosophers that take emotions to play a central role in grounding moral concepts (Smith 1976; Strawson 1962; Darwall 2013). On such accounts, anger is taken to implicitly make demands for acknowledgement of the harm done. Adam Smith writes that anger's aim is not so much to cause its target to 'feel pain in his turn... as to make him sensible, that the person whom he injured did not deserve to be treated in that manner' (1976: 95-6). Darwall (2013) takes the objects of anger to be 'also their addressees',

anger comes 'with an implicit RSVP' on whether the object of anger acknowledges that a harm has been done (89). Angry attitudes, therefore, do 'not seek to diminish, humiliate, insult, or damage their objects. The acknowledgment they seek is of a reciprocal standing' (89).

There is, therefore, a trend in moral philosophy to see one of anger's main aims as communicative, or epistemic, rather than punitive. Some might try to account for this epistemic aim within a retributive conception of anger. Recognizing that one has caused another harm can, after all, be a painful experience. Anger's desire might be for this subtle form of punishment. If this is so, then Aristotle might be right after all, in taking anger to aim for (some form of) revenge[18]. This is, I think, a misguided move, however, for the suffering of an agent can be an irrelevant, and even undesirable, outcome of the often merely epistemic aims for recognition of the harm done. In chapter 3 I defend an account of anger as aiming for recognition, where recognition is independent from retribution.

We have seen a move away from the traditional conception of anger involving a constitutive desire for retribution. Firstly, by advocating social constructivist theories of emotion, where constitutive ties to any desire is broken, and secondly, by highlighting a distinct desire linked to anger: that of retribution.

1.2.5 Taking Stock

In summary, we have identified four main points of disagreement between philosophers that have argued for the value of anger and those who have condemned the emotion. First, instead of viewing anger as opposed to rational thinking, anger is crucial to gaining epistemic insight for these thinkers. Second, although experiencing anger involves pain, it can also be a pleasurable experience, that has been argued to have important psychological benefits. These include resisting the internalization of oppression, allowing agents to affirm their worth, and increasing feelings of belonging when anger is shared.

Third, anger is taken to be effective, as opposed to counterproductive, in bringing about social change. Anger can be politically effective by motivating constructive non-aggressive

---

[18] See Christensen (2016), who reads Aristotle as being committed to a retributive construal of anger that includes such epistemic aims.

political actions, as well as by motivating aggressive action. Anger can be counterproductive in the short-term to angry agents seeking societal change, but this does not undermine its efficacy at bringing about long-term political change. Although those who praise anger do not deny that anger can manifest in aggressive and destructive actions, they do not think anger's political efficacy rests solely in its destructive potential.

Lastly, in place of endorsing a traditional view of anger where the emotion constitutively involves a drive for retribution, feminist philosophers have not been preoccupied with supplying universal definitions of the emotion. Despite this, there is a trend towards social constructivism regarding the emotions in the literature, which suggests opposition to the necessary connection between anger and retribution advocated by its critics. Furthermore, in so far as anger has been tied to an aim for recognition, and is cast as a crucial motivator for often constructive action, the link between anger and dispositions for punitive actions is weakened[19].

I have outlined four points of dispute regarding anger. I now turn to how we should tackle the debate surveyed.

1.3   Desiderata and Constraints

The debate on anger has not, so far, benefited from a foundational approach. By this I mean a 'bottom-up' approach that makes the emotion of anger central and asks questions of its metaphysics, epistemology and normativity. The literature has focused on questions of virtue, ethics and the consequences of anger, rather than on basic questions regarding anger itself. There seems to be space for a significant amount of foundational work regarding anger, work that might shed light on claims regarding its role in personal and political life. I propose that doing such foundational work will represent a novel and important way of moving the debate on anger forward. I outline four specific desiderata for moving the debate forward such as to map out some of the foundational questions in need of attention. These are, that

---

[19] It is important to note that commitments across these four domains need not line up neatly with each other. That is to say, particular views on anger may involve a mixture of commitments. One might think anger is psychologically beneficial despite being epistemically noxious, or vice versa, for example. One might think anger is beneficial to achieving one's practical ends while also believing it to be constitutively tied to retribution. Despite lacking necessary connections, the features associated with anger tend to come in the clusters surveyed above.

accounts be given of anger's: nature, instrumental rationality, normative rationality and epistemic rationality. I take the dispute over anger's practical utility in fighting against social injustice to relate to its instrumental (ir)rationality, while I take dispute over anger's role in rational thinking to relate to both questions of anger's normative, as well as, epistemic rationality. I will outline these desiderata below. In so doing, I also highlight two methodological constraints that will be crucial to fulfilling the desiderata outlined. These are, that we not abstract away from the social reality in which emotions occur, and that we pay attention to relevant experimental work from the brain and behavioural sciences. We will see that a foundational approach will lend support to those who praise anger.

The first desideratum concerns the provision of an account of anger's nature. Such an account is necessary, in the first instance, to assess the plausibility of a number of deflationary moves that could be made to dissolve the debate on anger. The deflationary moves I have in mind are ones that deny that those who condemn anger, and those who praise it, are in a dispute over the same phenomenon. We are, therefore, in need of a systematic account of anger's nature such as to assess whether there is a real debate at stake in the first place.

The thought is that the debate surveyed above might involve a mismatch in referent. Those condemning anger might be referring to instances of anger that are uncontrolled, vengeful and irrational, while those that take themselves to be opposed to this construal of anger are perhaps actually referring to a different type of anger; one that is rational, controlled and often productive. Joseph Butler, for example, endorses two distinct kinds of anger, one instinctive, 'hasty' and often irrational, the other a 'settled and deliberate' morally loaded resentment (see MacLachlan 2010)[20]. Perhaps we even have more than two types of anger at play, perhaps there are as many types of anger as there are combinations of the four features over which there seems to be dispute. Alternatively, we might think that anger in the social justice case is a particular, potentially even atypical, type of anger, such that what is true of anger in cases of social injustice is not true of anger in other cases. To address these possibilities, we need to consider anger's nature.

The anger debate has paid less attention to anger's nature than to its presumed effects. We saw that those who condemn anger have typically employed a retributive conception of the emotion, while those that have praised anger have either not been concerned with proposing

---

[20] Adam Smith has been read as endorsing two different types of anger, as he categories anger as an unsocial emotion in the first part of his work and later as an importantly social emotion, in the second part of his work. See MacLachlan (2010) for a critique of this view, however.

a positive alternative to this, or, have advocated a social constructivist theory of anger's nature, and highlighted its communicative rather than retributive aim as key. This leaves open the possibility that each side of the debate is actually referring to a different type of anger, or that one of the two sides is mistaken in thinking that they are discussing anger at all. We, therefore, seem to have reason to address the fundamental question of how best to construe anger. Doing so will help to better understand the nature of the debate and whether and where disagreement exists to begin with. Doing so will also help determine which side of the debate should be favoured. I, therefore, propose that the first desideratum for moving the debate forward is to give the nature of anger careful consideration.

Another reason to give anger's nature careful consideration, is that it will bear on the rationality of the emotion. If anger is a reflex-like state, it will likely lack intentional content. If this is the case, then anger will be best construed as an arational state. If, on the hand, anger is best thought of as an intentional state, it will have correctness conditions of some sort. We typically speak of emotions being 'appropriate' rather than correct. This suggests that emotions are states that are had for reasons. I use the term 'normative rationality' to refer to reason-responsiveness. For anger to be a state capable of normative rationality more will need to be said about the type of intentional state anger is. Furthermore, if anger is, by its nature, a modular phenomenon that constitutively involves particular action tendencies, then it risks being irrational by motivating actions that are triggered in isolation from, and often in opposition to, the agents broader intentions. I use the term 'instrumental rationality' to refer to anger's ability to further the emoting agent's practical aims. Careful consideration of the nature of anger will be crucial for shedding light on the normative and instrumental rationality of the emotion.

Anger's instrumental rationality concerns the consequentialist question of whether anger is beneficial to furthering the aims of the angry agent. Anger can cause many effects, behavioural, epistemic and psychological. Throughout, by 'instrumental rationality' I will mean anger's practical effects, namely the actions it motivates and the consequences of those actions. Anger's behavioural effects will, therefore, be central. I will say more about anger's epistemic effects below. Regarding anger's psychological effects, I will have little to add to what has been canvassed above. I take the first-hand accounts of oppressed individuals to speak to the complexity of anger's psychological effects. The emotion causes pain but also brings with it a powerful source of self-affirmation as well as identification with others. These psychological effects are bound to be critical in opposing the oppressive status quo, and

plausibly make anger somewhat pleasurable to feel. In light of these firs-personal accounts, I take claims that anger is psychologically instrumental irrationality, that is, that it undermines the angry agent's psychological well-being, to be misguided.

Regarding anger's instrumental rationality I will focus on whether anger motivates actions that further the aims of the oppressed in fighting against injustice. This is the positive instrumental role for anger that has been most forcefully advocated in the literature, and which has received the strongest recent critique. Our second desideratum is, therefore, to shed light on whether anger is instrumentally rational in this sense. I will specifically be concerned with whether anger is instrumentally ration in a manner that does not presuppose its link to destructive action tendencies. That anger can be effective at enacting social change, by means of violent revolt, is clear. Although we have seen that many contend that anger is instrumentally rational in constructively fighting against oppression, specific arguments to this effect remain lacking in the literature.

There seems to be a partly empirical question needing further investigation regarding anger, namely whether it's characterization as involving a desire to harm, and having the basic function of threatening, attacking, and protecting, are warranted. Such empirical questions require theoretical framing, and interpretation, and will have consequences for our view of anger's nature and instrumental rationality. This highlights the first of my methodological constraints for moving the debate forward, which is to pay attention to relevant and recent empirical work. As an embodied psychological phenomenon, any debate on anger will benefit from being empirically informed. The relevance of empirical work will be particularly important in meeting our first two desiderata; namely in investigating the nature of anger and the instrumental rationality of its effects.

Regarding anger's normative rationality, I will be concerned with whether, and how, anger is the sort of state that is reason-responsive. Emotions are often taken to be reason-responsive in contemporary moral psychology and philosophy of emotion, but elaborations on how this reason-responsiveness is to be construed are scarce (Raz 2011; Skorupski 2010; Deonna and Teroni 2012b; Tappolet 2016). Without stipulating otherwise, this leaves open the possibility that emotions are reason-responsive in much the same ways as are beliefs or intentions. Whether we should endorse this implication or not, and what the consequences of doing so are for the debate on anger, remains unexplored. If no feasible account of anger as reason-responsive is possible, then this lends support to those who would deny the emotions a place

within the realm of reason. If anger is reason-responsive in much the same ways as beliefs are, this might be seen as a victory for the defender of anger's knowledge producing role, but this might be too quick. If emotions such as anger are reason-responsive in much the same way as beliefs are, then they risk being epistemically dispensable. If this is the case, and some of their noxious epistemic effects are typically at play, then such an account of anger as reason-responsive could nonetheless lend support to those who condemn the emotion after all. My third desideratum is to provide an account of anger as reason-responsive. I will argue that emotions enjoy sui generis normative rationality, such that the ways in which anger is reason-responsive is irreducible to the reason-responsiveness of beliefs or intentions.

Once we have met this third desideratum, so that we have a working account of anger as reason-responsive on the table, we will be in a good position to consider anger's epistemic role. I use the term 'epistemic rationality' to refer to anger's epistemic role. This term is meant to capture anger's role in leading agent to form beliefs that are supported by adequate evidence. There are two main ways anger could do this: by acting as adequate evidence for beliefs itself, or by signaling to the agent that there is relevant evidence on which beliefs can then be based. The former would involve anger playing a *justificatory* epistemic role, while the latter would involve anger playing a *causal* epistemic role. The distinction here is between whether anger plays its putative epistemic role by causing agents to non-emotional acknowledge, or to search for, reasons for their anger, or instead, whether anger can itself directly justify the relevant evaluative beliefs. On both accounts anger is a guide to reasons, which would make anger a means to, as opposed to hinderance to, uncovering truths.

It is clear that for anger to be knowledge conducive, as its defenders claim the emotion is, then anger must be epistemically rational. What specific epistemic role is involved in anger's putative epistemic rationality is, however, not clearly stated, or argued for, by anger's defenders. We are, therefore, in need of elaboration on whether anger does indeed play a knowledge conducive role, and if it does, what sort of epistemic role this is.

Even if anger has a positive epistemic role to play, if anger typically causes the noxious epistemic effects as well, then anger might turn out to be epistemically irrational. That is, anger might undermine the ability of agents to form beliefs based on the relevant evidence despite also playing positive epistemic roles. To determine whether anger is epistemically rational, then, we might need to weigh its potential positive epistemic role against the noxious epistemic effects that anger allegedly causes. One can seek to diminish the weight of anger's

purported negative epistemic effects in two ways. First, by casting doubt on the evidence in its support, and second, by securing anger not only a strong epistemic role, but a crucial one.

We have seen that the feminist literature endorses the view that emotions play strong epistemic roles (Friedman 1986; Jaggar 1989), despite not providing precise accounts of the types of role these are meant to be. In addition to this, the literature claims that anger's epistemic role is often crucial, or necessary, specifically for gaining insight of social injustices under conditions of oppression. For anger's defenders to make a case for this, accounts of the specific epistemic role they envision for anger should be developed. The feasibility of their claims that anger indeed plays such roles will hinge on the particularities and strengths of the actual accounts put forth. The debate on anger will, therefore, be furthered by the provision of an account for how anger is meant to play its putatively strong, and distinctive, epistemic roles(s). As accounts to this effect remain merely programmatic in the literature, we have a four desideratum for moving the debate forward: developing an account of anger's epistemic rationality. If anger plays crucial epistemic roles, then the bar is raised for those seeking to condemn anger on account of outweighing detrimental epistemic effects.

The anger debate has developed largely in isolation from discussions in analytic philosophy of emotion. This is, I think, a hinderance to the debate. Using the resources of contemporary philosophy of emotion, we will be better placed to theorize both the nature and rationality of anger. The reverse will also prove true, however. Many of the social and political features of the anger debate just outlined will help inform analytic thinking on anger, which has largely proceeded in abstraction from such concerns. This gives us our second methodological constraint: that we do not abstract away from the social and political dimensions of anger in the course of this project. To do so would be to fail to contribute to the debate on human anger. To do so would also, I think, result in worse theorizing.

The points of disagreement between the two stances on anger surveyed above, all relate, in some way, to assessments of anger's rationality. Whether anger is counterproductive to the aims of the angry agent relates to what I call anger's instrumental rationality, in furthering, or undermining, the agent's aims. Whether anger is psychologically beneficial to the emoting agent relates to whether anger is instrumentally rational in a distinct sense, namely whether it is a positive state to experience with regards to its psychological effects. Whether anger is an aid, or a hinderance, to epistemic life relates both to what I call anger's normative rationality, that is, whether anger is a state capable of responding to reasons, as well as to

what I call anger's epistemic rationality, that is, whether anger has a positive or negative effect on leading agents to form beliefs based on adequate evidence. Lastly, anger's nature is related to its rationality, albeit indirectly, for our theory of anger will impact what types of rationality the emotion is capable of, as well as have implications for the form its distinct types of rationality might take. I have said that there is enough to deny anger is instrumentally irrational on account of its psychological effects, given that its effects in this domain are mixed in most cases, and particularly positive in those cases where anger is experienced by those living under conditions of oppression. I have outlined four desiderata for tackling the remaining senses of anger's rationality. These four desiderata can, therefore, be seen as joint desiderata for the provision of an account of the rationality of anger.

1.4   Conclusion

I began by characterizing two opposing views of anger, one proposed by those who condemn the emotion, and the other by those who praise it. I highlighted four points of substantive disagreement between the two camps. These were over; (i) whether anger is a benefit or a hindrance to rational thought, (ii) whether anger is practically beneficial to the angry agent, or whether it is most often counterproductive to securing their aims, (iii) whether anger is promoting or diminishing of the psychological well-being of the agent and, lastly, (iv) whether anger is by its nature tied to harming others or retribution.

I have said that I will take anger to have complex psychological effects that undermine any attempt to advocate against anger on their account. The testimony of victims of gender and racial oppression are sufficient to establish the therapeutic and emancipatory psychological effects of anger. The other three points of disagreement will be my concern. All of them relate to the question of anger's rationality. These three points of disagreement generated four desiderata for moving the debate forward. These were:

1. The provision of an account of anger's nature
2. The provision of an account of anger's instrumental rationality
3. The provision of an account of anger's normative rationality
4. The provision of an account of anger's epistemic rationality

We need a systematic, rather than formulaic, account of anger's nature for a number of reasons. First, to assess the plausibility of deflationary attempts at dissolving the debate by appeal to a mismatch in referents. Second, to determine whether the traditional view of anger as constitutively involving a desire for retribution is a plausible one. Third, to shed light on questions of anger's rationality.

To assess whether anger typically undermines the aims of emoting agents we need an account of anger's instrumental rationality, or lack thereof. We need an account of what sorts of aims we have in mind, and how anger tends to further, or undermine, them. I focus on the case of social justice, where the aim of angry agents is to improve their current social situation. I focus on whether anger can be instrumentally rational in achieving this aim, without being counterproductive, in the short-term, for the angry individual. That anger can prove instrumentally rational in fighting against social injustice is not contentious, whether it can do so in a manner that does not typically undermine the aims of the angry agent, at least in the short-term, is unclear.

We saw that those who praise anger disagree with those who condemn it over whether the emotion's epistemic value is positive. Those who condemn anger seem to take anger to either be an arational state, such as a bodily disturbance that disrupts rational thinking, or a reason-responsive state that tends to lead the agent's reasoning astray. Those who praise anger, on the other hand, take anger to be a reason-responsive state that plays positive epistemic roles in producing knowledge. They do not, however, provide specific accounts of anger's reason-responsiveness or epistemic role. My third and fourth desiderata are complimentary in that they relate to the epistemic value of anger. To assess whether anger's epistemic role is positive, we need an account of anger's normative rationality, that is, of whether and how anger is a reason-responsive state, as well as an account of anger's epistemic rationality, that is, whether and how anger plays a role in the formation of justified beliefs.

Meeting the first two desiderata, in particular, will benefit from attention to recent experimental work in the brain and behavioural sciences. That attention be paid to such work was my first methodological constraint. As a crucial epistemic role for anger has been stressed under conditions of oppression, meeting the third and fourth desiderata will particularly benefit from attention to the oppressive relations that structure social reality. That my project remains sensitive to facts pertaining to social reality was my second methodological constraint.

Pursuing these desiderata will allow me to provide a, currently lacking, more foundational approach to the debate on anger. Doing so will reveal anger to be a phenomenon capable of all three types of rationality outlined. A foundational approach to the anger debate will strengthen the case of those who praise anger, by providing a systematic account of anger as rational. The positive view of anger has been disproportionately neglected, and holds reduced currency, in folk psychology as well as analytic academic philosophy. My thesis takes steps to remedy this. Anger will emerge as a state that is not constitutively tied to retribution and that is instrumentally rational in fighting against social injustice. Crucially, we will see that it can be instrumentally rational without proving counterproductive to the angry agent in the short-term. We will see that anger is a state governed by sui generis reason-responsiveness, and that this in turn allows anger to play distinctive and crucial epistemic roles.

2

What is Anger?

*'Philosophers suppose that they are uncovering the true nature of emotion
as revealed a priori in vernacular emotion concepts...Using these concepts as a guide
to the emotions is like studying female sexuality by reading pornography.'*
-Griffiths, in 'What Emotions Really Are'

What is anger? Responses to questions regarding the nature of emotions typically have the joint desiderata of supplying an answer that both matches sufficiently well the extension of our common linguistic use of the concept, and is projectable, or theoretically useful, within the confines of a particular discipline (DeLancey 2002). One of the most influential responses to this question, provided in the scientific study of emotion, is that some emotions are, or constitutively involve, basic affect programs. This notion of 'basic' emotions plays a central role in the philosophical literature on emotions (Robinson 2005: 33; Deonna and Teroni 2012: 19; Colombetti 2013: 26).

The idea of 'basic emotions' owes its popularity in large part to perhaps the most influential theory in affective science: basic emotion theory. According to basic emotion theory, a limited number of emotions are associated with a set of regular and hardwired features that have been called 'affect programs' (Ekman 1982; Griffiths 1997)[21]. Each of these basic emotions has its own affect program characterized by: quick automatic appraisal of distinctive antecedent events, distinctive universal facial expressions and physiological changes, distinctive subjective experience, and distinctive motor outputs (Ekman 1999: 56). Although the list of basic emotions varies, anger invariably makes the list (Ekman 1982; Williams 2017)[22]. Anger is thought to involve appraising an object as blameworthy for an undesirable situation, typically an offence or injury. When 'anger erupts, blood pressure will rise, a scowl will form on the face, and there will be an urge to hit' (Barrett 2006: 33). Affect programs are quick, short-lived and modular[23] (Deonna and Teroni 2012: 20).

Here I am following others in using 'modular' in the Fodorian sense (Griffiths 1997; Bluhm 2006). Specifically, the claim is that affect programs share at least four of the nine features Fodor characterized modular systems as possessing[24]. Affect programs are fast, associated with specific neural circuits, mandatory (i.e. automatic or outside complete conscious control), and informationally encapsulated or cognitively impenetrable, such that the processing of modules is insensitive to the agent's beliefs, desires and intentions[25].

Affect programs are seen as functional specializations that arose in response to adaptive pressures (Griffiths 1997). It is therefore the adaptive relevance of sets of stimuli, and the adaptive benefit of certain behaviours in response to them, that give rise to the distinct

---

[21] This does not mean that basic emotions are necessarily equated to their affect programs, but that those emotions that involve affect programs are basic. An emotion 'involving' an affect program is compatible with not all components of the relevant affect program manifesting in a particular instance of a given emotion. That one does not always furrow one's brow in anger does not mean anger does not involve an affect program that typically causes this facial expression.

[22] Other canonical basic emotions include disgust, fear, happiness, sadness, and surprise (Ekman 1999).

[23] There are many senses of 'modularity', see Faucher and Tappolet (2006) for a review of the different senses of modularity regarding the emotions.

[24] The other five features of Fodorian modules are: domain-specificity, limited central accessibility, 'shallowness' of output, characteristic and specific breakdown patterns, and characteristic pace and sequencing of ontogeny (Fodor 1983; Faucher and Tappolet 2006). Anger might fulfill all nine criteria, given that many of these are closely related features, but I highlight the four mentioned above as these have been those most commonly invoked in the characterization of affect programs.

[25] Note that informational encapsulation is not incompatible with input from higher-level processing, what this feature describes is that the top-down processing involved in any Fodorian module is strictly restricted to a specific set or 'database'. That it is to say, some beliefs and intentions might be relevant to modular processing, but not any. This might mean, for example, that the anger affect program is not entirely encapsulated from memories of your past offences, but that the module's processing is encapsulated from your wider beliefs about the object of your anger, as well as from desires not to become angered.

modular affect programs that characterize basic emotions. On basic emotion theory then, basic emotions are those emotions which involve evolved affect programs.

If anger involves an affect program of the sort described, this could lend support to those who condemn the emotion. We saw that those who condemn anger tended to characterize the emotion as immune to reason, as instinctive, and animal-like. If anger is merely a reflex-like state, it might be best construed as arational and therefore outside the realm of reason. If anger is a hardwired to respond to stimuli from our distant evolutionary past, it might prove maladaptive to the modern day. Furthermore, if anger involves a modular affect program that served the adaptive function of protection against threats and harms, initiating attack behaviour in response to these, it might typically lead to irrational action. A basic construal of anger then gives us a picture of anger quite in line with that of those who condemn it[26]. That is, basic anger seems like an adequate target for condemnation, given that it is triggered independently to one's wider concerns, and that it typically causes hardwired behaviour, in a manner encapsulated from higher-level reasoning. There is therefore reason to take a considered look at what basic emotion theory is actually committed to regarding anger, and whether its commitments hold up to scrutiny.

It would be commonplace to point out that not all instances of human anger match the basic emotion theory's characterization. We don't invariably, or perhaps even typically, have an urge to hit people with whom we are angry in our everyday lives, nor do we always scowl in rage. While anger is a common everyday experience for ordinary individuals, engaging in physical violence when in anger, typically, is not. Basic emotion theory, as we will see, can account for this by holding that the outputs of affect programs can be significantly molded by upbringing, such that hardwired behavioural tendencies are suppressed. Proponents of basic emotion theory can also appeal to affect programs being 'prototypical' of basic emotions, such that every instance of anger need not involve all features of its affect program to count as an instance of basic anger.

---

[26] I do not mean to suggest that those who condemn anger must endorse a basic emotion theory of anger. One could of course argue that anger is condemnable while upholding a social constructivist view of anger. The prevalence of a basic emotion construal of anger in the philosophical literature is, however, widespread (Griffiths 1997; Deonna and Teroni 2012; Nussbaum 2016) and I take it to significantly contribute to the prevalence of a simplistic view of anger, as well as to its bad reputation. Indeed, it has been noted that the modular view of emotions is very much in line with construing them as prima facie opponents to reason and rational behaviour (Faucher and Tappolet 2006).

Basic emotion theory allows for the existence of instances of anger that depart significantly from the anger affect program however. These would be instances of non-basic anger, which are phylogenetically younger than, and evolutionarily dependent on, basic anger. Non-basic anger is less modular and involves heavier input from cognitive or higher-level processing than non-basic anger. It is also more culturally variable and typically not shared by other animals (Griffiths 1997).

This suggests that there might be two different kinds of anger that our common use of the emotion term does not discriminate between: basic affect program types of anger, and non-basic, more cognitive, types of anger. This is indeed the extension of the folk concept of 'anger' as Griffiths (1997) sees it. Here the option of there being confused referents emerges, where those who condemn anger might have basic anger in mind, while proponents of moral anger might instead be referring to non-basic anger. One way of dissolving the dispute about anger would be to distinguish between two types of anger, and to condemn one while accepting the other as valuable. An alternative would be to argue that 'anger' only properly refers to one of the two phenomena, thereby rejecting the claim that one's opponent's view should be taken seriously as making claims regarding anger at all.

A close look at basic emotion theory will speak against all of the options just mentioned. Attention to work in evolutionary biology and psychology will reveal the anger affect program to be hard to specify, it's modularity questionable, and the distinction between basic and non-basic forms of anger hard to uphold. This will lead us to endorse a heterogenous view of anger, where our common usage of the emotion term refers to members of the anger family that share an affect program ancestor. We will see that anger plays multiple evolutionary functions and that there is no robust distinction between basic and non-basic anger.

There are three main related, but distinct, potential claims made by basic emotion theory in holding anger to be a basic emotion;

1. That anger is *universal*, that it is shared across all cultures and has similar expressions in animals, particularly non-human primates.
2. That anger is *biologically basic*, hardwired due to a beneficial function in our evolutionary history.

3. That anger is *psychologically basic* such that it doesn't have other emotions as constituent parts.

I will deal with each of these in turn and argue that only a very thin notion of anger as a basic emotion survives scrutiny. This will in turn lend some support to a departure from the traditional conception of anger as serving a basic attack related function. After my critique of basic emotion theory, I defend my view of anger against the possible charge that it does not form a natural kind. Doing so will allow me to dismiss any worries pertaining to the reliance on empirical work based on such concerns. I finish by situating my construal of anger in relation to existing theories in the philosophy of emotion literature by highlighting a set of uncontroversial commitments with which I am aligned.

2.1   Anger as Universal

Ekman and Izard famously extended Darwin's contention that human emotions share similar cross-cultural expression (Ekman 1982; Izard 1994). Amongst others, they investigated the facial expression of emotions across Europe, the US, Japan, and Africa. They found similar consistent results cross-culturally: participants could successfully match the emotion terms – anger, disgust, sadness, happiness, fear and surprise – to characteristic facial expressions. This was found even in cultures that had experienced virtually no contact with western society, such as the Dani of Iran, and the South Fore of New Guinea (Ekman and Friesen 1971). This led Ekman to conclude that these six emotions are innate and universal. Although there is dispute amongst basic emotion theorists about which emotions count as basic, anger invariably makes the list (Ortony and Turner 1990).

Basic emotion theory has come under attack on a number of fronts[27]. For one, the experimental evidence levied in support of the theory is vulnerable to critique. I will mention only a few methodological concerns. Firstly, results were gathered by asking participants to choose an emotion from a prespecified list of emotions. A forced-choice experimental design of this sort cues the subject to the experimenter's expectations. Different results have actually been reported when the prespecified emotion terms on the list are changed, confirming the

---

[27] For psychological critiques of the basic emotion theory see Ortony & Turner (1990) and Russell (1994). Philosophical critiques can be found in Colombetti (2013: 29-46), Deonna and Teroni (2012: 18-26), and Bluhm (2006). For philosophical defenses of basic emotion theory see DeLancey (2002) and Scarantino and Griffiths (2011).

significance of the options available to the subject in skewing results (Russell 1994). Furthermore, the experimental paradigms used were linguistically heavy and involved translation of questions into languages other than English as well as translation of responses back into English, weakening the robustness of the findings. A final methodological worry centers on the fact that the images of facial expressions used were posed or acted out, rather than spontaneous genuine emotional expressions. This further weakens extrapolations based on experimental observations to real life emotions (Plamper 2017: 159).

Facial expressions cannot be established as necessary or sufficient for human emotions as we know that we can feel sad or angry without displaying the stereotypical facial expressions associated with each respective emotion. The universality of these expressions is proposed as prototypical, rather than definitional, of emotions. Key anthropological work is taken to present a challenge to the universality of prototypical emotional expressions, as psychologists are accused of having drastically misconstrued, or even ignored, ethnographic evidence (Solomon 1984; Deonna and Teroni 2012a). I briefly review the anthropological work on anger, such that we can consider what sort of challenge it poses against basic emotion theory.

Two societies are claimed to be void of expressions of anger; The Tahitians, and the Utku (Utkuhikhalingmiut) Eskimos (Briggs 1971; Levy 1973). If this is so, this would challenge the universality thesis of basic emotion theorists, at least regarding emotion expression. The Utku Eskimos, as the title of Briggs (1971) work *Never in Anger* foreshadows, do not seem to express anger. Although Briggs takes them to experience anger, she observes the remarkable suppression of the expression of this emotion amongst the Utku. They do not talk about anger, nor do they have a set of words to designate the emotion (Briggs 1971; Solomon 1984). Anger arguably takes on a more powerful role in Utku life than in western society, for on the rare occasions in which it is expressed, it is to be taken seriously. Indeed, an Utku reportedly, in speaking about Briggs, said that if Briggs got angry one should not pay her much attention, and just 'leave her alone', on the other hand 'if an Eskimo gets angry it's something to remember… a *kapluna* (white person) can get angry in the morning and be over it in the afternoon' (Briggs 1971: 261).

The Tahitians for their part, according to Levy (1973), place a lot of importance on anger (*riri*). Instead of the emotional control of the Utku however, the Tahitians verbally expressed their anger without constraint. They talked about anger often, and 'hypercognized' it

extensively, having around 50 words for referring to it (Levy 1973; Solomon 1984). The Tahitians believed that anger 'has bad effects on one's body' (Levy 1973: 285) and sought to quickly rid themselves of anger through dialogue and action. 'These are gentle people' Levy (1973: 285) writes, and there are 'extremely few reports of angry behavior'. Morrison wrote, two centuries before Levy, that the Tahitians are 'slow to anger and soon appeased' (275), and Levy cites a policeman that observed the Tahitian's 'lack of a vengeful spirit' (276).

Another famous piece of emotion ethnography, conducted by Catherine Lutz (1988), identifies the emotion of *song* in the Ifaluk people. In contrast to the western notion of anger, *song* is limited to instances where one is justifiably angry, and involves no outbursts or loss of control, with which the western notion is often associated (Plamper 2017: 107). Lutz writes that 'daily negotiations over who is *song* and over the proper reasons for that anger lie at the heart of the politics of everyday life' (1988: 170) and 'it is expected that those who are justifiably angry [*song*] will not physically aggress against another. And in fact, interpersonal violence is virtually nonexistent on the island' (176).

These ethnographic accounts of anger are often levied as evidence against the universality claim of basic emotion theory. The thought is that if the expression of anger is drastically different in these communities, then the anger affect program is not *expressed* universally or cross-culturally, as basic emotion theorists claim. This, in turn, might be evidence of something stronger, namely that the anger affect program *itself* is not present across cultures. The former claim suggests that culture is able to mold expressions quite heavily, but does not challenge the existence of affect programs themselves. The latter claim would undermine the universality of affect programs themselves, rather than merely the universality of their expressions. A claim of this latter sort is assented to by social constructivist theories of emotion, where emotions of different cultural (and historical) communities are taken to *be* and *feel* fundamentally different.

It is important to note that neither Briggs nor Levy subscribed to a constructivist account of emotions - their accounts focused on expression rather than feeling[28]. They, would, therefore likely not be committed to the stronger claim that emotions have a socially constructed nature, but merely to the claim that emotional expression is not universal (Plamper 2017:

---

[28] Briggs and Levy wrote before the social constructivist turn in anthropology and did not deviate from the following disciplinary trend of their time: that 'emotions have a uniform core but are expressed variously in different cultures- this summarizes the guiding assumption of the anthropology of emotions during the 1970s' (Plamper 2017: 98).

95). Lutz, on the other hand, is a social constructivist about emotions, she writes that 'emotional experience is not precultural but preeminently cultural' (1988: 5). Social constructivist views of emotion take the rich, multi-faceted cultural context in which emotions unfold to be constitutive of emotions[29], such that anger might be experienced and expressed variably across, or even absent from, cultures in which it takes on different social functions, and meanings, to our own.

The latter claim, that affect programs themselves are not universal, cannot be established by the above-mentioned ethnographic research. Emotions themselves cannot be observed, only their expressions can. It is, therefore, unlikely that it can be known whether the Utka, or the Tahitians, experience anger or not, and similarly whether the *song* of the Ifaluk feels how our everyday anger feels. It is clear that anger plays a different role in these societies and I am sympathetic to the view that anger might indeed be experienced in significantly different ways to our own, but the evidence supplied in such ethnographies cannot establish this claim. How can the anthropologist determine which emotions are experienced (or not) other than by assuming the universality theory that their work is levied as evidence against? If the anthropologist takes the lack of what they take to be typical expressions of anger, to be indicative of the absence of the emotion our western folk concept of anger refers to, they seem to be assuming the very theory their work is levied as evidence against. That is, they assume anger can be identified by a set of distinctive expressions, and therefore that there is a stable emotion type that these expressions are expressions of[30]. Solomon (1984) has made a similar point and writes:

> Let me raise the problem, as philosophers are prone to do, to the level of an explicit paradox: the thesis or hypothesis under consideration is the alleged difference between Utku emotion and our own, particularly regarding emotions like anger. But

---

[29] There are many ways of cashing out social constructivist claims, that differ depending on the target of construction. At the trivial level, any object that depends on humans for their existence is socially constructed (cars for example). The philosophically interesting conceptions of the social construction of objects have revolved around the notion of *constitutive* social construction, as applied to objects that appear (in the first instance) not to be based on cultural mechanisms: such as sex differences, emotions, and race (Mallon 2007; Marques 2015). The thought is, therefore, that emotions are socially constructed if some of their constitutive features are created or sustained by social and cultural practices. Lutz, for example, takes anger to be and feel fundamentally different, depending on cultural differences in meaning and practice. The kind 'anger' is therefore distinct across cultures. The main form social constructivism of emotions in psychology is the psychological constructivism of Barrett (2006) and Russell (2003), where emotions result from the categorization of felt bodily states into folk emotion concepts. The culturally dependent *concepts* for different emotion types are therefore constitutive of emotional experience on such accounts. See Kurth (2018) for a critique of the psychological constructivist view.

[30] There is also the meta-linguistic worry, which has plagued my previous paragraphs, that it becomes difficult (or impossible) to talk of *what* is being socially constructed if we do not assume a basic common referent. The anthropology of emotion is plagued by this worry, in so far as it seeks to establish some object X is socially constructed in a strong sense. Hacking (1999) makes this point.

> insofar as the anthropologist assumes that she is capable of understanding the emotional expression of her subject, that is, understanding them as expressions of particular emotions then she must assume from the outset precisely the hypothesis to be verified, namely, that different people have, and can mutually understand, essentially the same emotions. (84)

The ethnographic work, therefore, cannot establish that anger is felt differently in different communities. What the ethnographic work does and should draw our attention to however, is the various ways in which anger manifests itself, how subjectable to cultural influence it is and how it can take on drastically different meanings in different societies.

Despite the methodological problems just highlighted, often the literature seems to portray a fierce debate about the universality of emotions, with basic emotion theorists on one side, and anthropologists and historians on another (Solomon 1984; Mallon and Stich 2000; Colombetti 2013). The debate purports to hinge around whether emotions form 'natural kinds', with basic emotion theorists holding that basic emotions are natural kinds, while anthropologists oppose this view, given their endorsement of the heterogeneity of affective phenomena across cultures[31].

The conflict between the two can be minimized however. Basic emotion theorists have themselves claimed that the causes of emotions, as well as their expressions, can be highly variable, and that they can be learnt, making them culturally dependent. What they claim is that the affect program, i.e. the prototypical expressions of the emotion, are nonetheless universal (Ekman 1982). On such a view, children are reared to associate certain stimuli with a particular type of emotional response and the expression of the emotion is influenced by context specific display rules that suppress or otherwise modify the hardwired expressions of that emotion type. This is what Goldie (2002) calls the 'avocado pear' picture of emotions, where one thinks of the emotions as 'comprising, under the surface skin, a soft outer structure (that which varies culturally), and a hard inner core (that which is biological and universal)' (99).

Another way of diluting the conflict is to construe the disagreement between basic emotion theorists and anthropologists as one about what emotion terms actually refer to. The meaning and reference of 'anger' for the basic emotion psychologist is just the stereotypical features associated with the emotion, while for the anthropologist, the meaning and reference

---

[31] I say more on natural kinds in section 2.5.

of 'anger' is something thicker, and involves the culturally specific dimensions of the emotion, such as display rules and contextual meanings. The two views then, need not be opposed. Mallon and Stich (2000) have developed an argument of this sort. Figure 1 summarizes the distinct referents they take to be at play.

Antecedent Conditions

Recruited Response Tendencies

Measurable responses

Interpersonal

Emotion Prototype

Subjective Experience

Self-Report

Facial Program

Facial Expression

Vocalization Program

Voice Tone

Intrapersonal

Motor Program

Motor Behavior

Physiologic Support

Physiologic response

Appraisal

Cultural learning: Display and feeling rules

The referent of 'fear' for Social Constructionists

The referent of 'fear' for Evolutionary Psychologists

Figure 1: Two different emotion referents for 'fear' (Mallon and Stich, 2000)

These attempts at consolidating basic emotion theory and the ethnographic research subscribe to a traditional conception of biology/culture divide. Both the basic emotion theory's attempt to account for variety, by allowing a significant role for learning, as well as the offering of a mismatching referent approach, endorse an outdated conception of a dichotomy between the biological and the social. The former says that the biology is the same, but inputs and outputs are influenced by culture, while the latter says that there is a semantic disagreement about whether to include the peripheral cultural effects in one's conception of emotion, or limit one's concept to the biological 'core'.

These are not the only options on the table. The past three decades have seen a remarkable increase in work that blurs the distinction between biology and culture. From epigenetics, and the social dependency of the development of many traits, to the ever more popular

concept of cultural evolution, this distinction is losing ground[32]. These are all forms of exogenetic inheritance (Griffiths and Stotz 2013), where changes in phenotype do not involve a corresponding change in genotype. Epigenetics highlights the influence of extra-genetic environmental features (such as temperature, age, lifestyle) on gene expression (Jablonka and Lamb 2006), while developmental biology has highlighted the importance of social interaction for the proper functioning of mature physical systems. Griffiths (1997) for example writes that:

> the social interactions that induce normal psychosocial development in the rhesus monkey are as much part of its developmental system as the endoplasmic reticulum[33] of its maternal gamete. (128)

There is evidence of the social dependence of emotional development as well. The reciprocal connections between pre-frontal cortex and the limbic brain areas associated with emotions, for example, are developed and maintained by mother-infant interaction in early life (see Bluhm 2006). This highlights the importance of social interaction for the development of emotional phenotypes.

Cultural evolution further expands the domain of things that can be seen as evolved beyond the biological, where inheritance is not the purview of genetics alone (Boyd and Richerson 1985; Sober 1991; Sterelny 2006; Jablonka and Lamb 2006). As long as a feature is reliably self-replicating, then it can be inherited, for example:

> If a child-care practice is stably transmitted in a lineage and affects the psychological phenotype, then its invention constitutes microevolution as surely as a change in allele at a locus (Griffiths 1997: 131).

Similarly, in the emotional realm, provided specific interpersonal emotional practices were stably transmitted over time, these emotional expressions and practices 'evolved' – resulting in an emotion that is, in part, the product of cultural evolution.

I follow Griffiths (1997) in taking emotions to be 'heterogeneously constructed'. They aren't mostly biologically determined, with then learnt cultural variations taking place on the

---

[32] Epigenetics, in the narrow sense, refers to the influence of environmental factors in determining which genome sequences will be expressed in a cell. In a broad sense, epigenetics refers to any exogenetic inheritance, that is, to any causal pathways that influence the inheritance of phenotypes without changes to genotype (Griffiths and Stotz 2013). The social dependency of development, as well as cultural evolution, can be seen as forms of epigenetic inheritance, construed in this broad sense. I follow Griffiths and Stotz (2013) in using the term 'exogenetic' for any form of inheritance that results in phenotypes without changes to genotype, and reserve the term 'epigenetics' for its narrow usage, within organisms.

[33] The endoplasmic reticulum is the main organelle responsible for protein synthesis.

timescale of child development. They also aren't wholly determined by our social practices and folk concepts. Emotions are heterogeneously constructed in that they are individuated by evolved biological traits which are subject to a range of exogenetic evolution, including: epigenetic inheritance, developmental dependence on social interaction, and cultural evolution[34]. This claim does not only apply to modern emotions. Reliably self-replicating cultural practices, epigenetic influences and socially dependent development, have been at play long before the emergence of homo sapiens, and such exogenetically inherited features have contributed to the nature of modern-day emotions[35].

What this suggests is that anger is not going to turn out to be universal in quite the sense intended by basic emotion theory. There are unlikely to be modular affect programs for anger that have remained unchanged by shifting adaptive environments, and changing social and cultural practices, throughout our evolutionary history. Is there nothing then that all instances of anger have in common? I propose that we take anger to be universal in the following minimal sense: Anger is universal in that it is a culturally variable kind that is present in some form across our species' population. That this be the case suggests that what all these anger phenomena have in common is, not that they 'involve' a common modular affect-program, but that they all have an anger affect program *ancestor*. This basic anger ancestor is one we most likely share with other non-human primates, and which explains the proximity in some of our emotional responses. What is basic about anger then, is that there is a common ancestor to the variety of anger phenomena witnessed across the human population.

This ancestral anger affect program might be characterized by a particular function, perhaps an urge to attack, but that this be the case does not give us reason to characterize the function of human anger in line with that of a distant common ancestor more than it gives us reason to characterize the function of arms in terms of the function of fins. We will shortly see, however, that it is not easy to determine a basic evolved function for anger to begin with.

---

[34] As well as by factors that are hard to categorize in terms of this dichotomy (Griffiths 1997: 132).
[35] Here I am understanding 'culture' along the lines as those who advocate for cultural evolution. 'Culture' involves any 'information capable of affecting individuals' behavior that they acquire from other members of their species through teaching, imitation, and other forms of social learning' (Richerson and Boyd 2005: 5).

2.2   Anger as Biologically Basic

The most popular understanding of 'basic' is an evolutionary or biological one, from which the other two construals might be argued to follow. If anger is a biologically basic phenomena, then it is assumed that it must have served some evolutionary function in ensuring survival, that led it to be selected and survive into modern times. The evolutionary benefit of anger, much like other that of other emotions, is taken as given in the philosophy of emotion literature.

> Affect programs are an adaptive response to events that have a particular ecological significance for the organism. (Griffiths: 1997: 89)

> Anger has a very limited but real utility, which derives, very likely, from its evolutionary role as a 'fight-or-flight' mechanism. (Nussbaum 2015: 55)

> The social effect of expressed arousal may well have been a selector in the pre-human and early human evolution of certain emotional capacities (alarm and fierce anger are obvious candidates). (Pugmire 2005: 86)

What selective advantage did anger actually confer though? Those who condemn anger are likely to think that anger's function in seeking retribution was evolutionarily selected for its adaptive benefit. Indeed, Nussbaum (2016) says of anger that 'the desire for payback probably expresses an innate evolutionary tendency' (34). Although the idea that anger must have fulfilled some beneficial function is ubiquitous, what this function was is all from clear. Proposed functions have included; deterrence of harm (Tooby and Cosmides 2008), harming offenders (Fessler and Haley 2003), ensuring individual benefit (Sell, Tooby, and Cosmides 2009), enforcing communal norms, including fairness (Fessler and Haley 2003; de Waal 1997), communication (Ekman 1999; Pugmire 2005) and the promotion of cooperation (Frank 1988).

These functions are not mutually exclusive, and perhaps more than one of the above gets things right, depending on the level of analysis. Functional accounts of any phenomena tend to vary depending on the level of analysis. I will focus on three levels that have been distinguished for the functional analysis of emotions; individual, dyadic, and group level (Keltner and Haidt 1999). The first is concerned with the functions of emotions at the individual level. These are typically twofold: gaining important information from the environment, and responding in a behaviorally appropriate way to the emotion elicitor. At

the individual level then, the function of anger is to draw the individual's attention to objects that block or frustrate the agent's aims, as well as to (typically intentional) harms committed against them by others (Keltner and Haidt, 1999). In addition to this, anger functions at the individual level to prepare the agent to act in response to these stimuli. Anger therefore has both *epistemic* and *behavioural* functions at the individual level of analysis.

The dyadic level of analysis is concerned with the effects of emotion on others, within meaningful relationships broadly construed (Keltner and Haidt, 1999). This includes the function of anger in communicating the individual's emotions, beliefs and intentions. Different types of information can be communicated, however. The observer might come to share the emoting individual's emotion and, therefore, plausibly share in relevant beliefs about the state of affairs. For example, in observing an angry individual one can come to share their outrage at the current state of affairs. This spreading of an emotional state to others can be very helpful in facilitating cooperative behaviour. Communicating anger can also cause fear however, as the anger might communicate an intent to harm, and therefore act as a deterrent against future harms. At the dyadic level of analysis we, therefore, find at least two distinct functions for anger; *promoting cooperative behaviour* by recruiting others to share one's emotion or simply by communicating one's intentions and beliefs, as well as *deterring* harmful, or otherwise undesired, behaviour by communicating one's commitment to self-defense (Ekman, 1982; de Waal 1997; Fessler and Haley, 2003; Tooby and Cosmides 2008). These functions would have adaptive benefit by ensuring others respect you and do not repeat their offence, as well as by facilitating social cooperation, which early hominid survival was heavily dependent upon (de Waal 1997; Brosnan 2006).

At the group level of analysis, emotions play roles that promote group goals and help regulate group relations. Emotions occurring in group contexts may confer a number of advantages: by facilitating *cooperative group behaviour* and by *enforcing communal norms* that regulate communal life de (Waal 1997; Fessler and Haley, 2003; Brosnan 2006; Tooby and Cosmides 2008). Anger in particular has played important roles in both group level functions of emotions. Collective anger is a clear motivator of action against another group, and anger's communicative function can quickly recruit in-group members to share in rage, and facilitate acting against offences collectively. Anger is also thought to play a role in maintaining social order, both in maintaining and enforcing hierarchical social relations, and in punishing group-members for norm violations. When an individual behaves in a way that doesn't

favour group cooperation, other group-members often express anger as means to enforce norm compliance. Non-human primates have been observed to punish freeloaders in aid of cooperative efforts, for example (Brosnan 2006; Suchak et al. 2016). Non-human primates also demonstrate inequity aversion: they become enraged by the unfair distribution of goods amongst conspecifics (de Waal 1992; Brosnan 2006). The more cooperative a species is, the higher the rates of inequity aversion (Brosnan 2006). Chimpanzees, for example, show not only angry responses when given less food than a con-specific, but also show aversion to receiving more than a con-specific, and will, therefore, choose fairer distributions even at a cost to themselves (de Waal 2014). Aversion to inequality is taken to be a mark of highly cooperative species such as our own. Inequity aversion in non-human primates is likely a homologue of human anger at social injustices. Anger, therefore, plays distinctively cooperative functions, as can be observed in our primate cousins. We, therefore, have two further adaptive functions that anger may be playing at the group level of analysis[36].

These three levels of analysis, the individual, dyadic and group levels, all make claims about the evolutionary benefit conferred by anger. The specific evolutionary benefit conferred differs however, depending on the level analysis, and there is more than one function for anger proposed at each level. There are a few options for where to go from here. We can argue that anger has one basic function, from which the others arose, perhaps through further adaptation, or we can accept that anger has a number of functions, none of which is more primary than the others. A case might be made for the former with regards to anger's role in facilitating cooperation and enforcing group norms. As we evolved into ever more social animals, the basic function of anger to register and act on harms committed against oneself, at the individual level, may have been adapted to become triggered by norm violations that do not involve oneself, and to be triggered by norms that violate cooperation such as unequal distribution.

If we endorse an evolutionary story such as this one, and our interest is human anger, it becomes misplaced to prioritize as primary the phylogenetically more ancient function. To

---

[36] How to make sense of group selection is a topic of controversy in the philosophy of biology (Okasha 2001; Godfrey-Smith 2008; Okasha and Paternotte 2012; Sterelny 2016). This controversy need not preoccupy us however, for it revolves around whether the concept of adaptation is best applied to groups in the explanation of evolved phenomena. There are a number of proposals in the mentioned resources in favour of a robust sense of group adaptation, but even if adaptation remains the purview of individuals alone, the cooperative and norm-enforcing functions of anger highlighted would remain adaptive functions, just ones that can be made sense of without a technical sense of 'group selection'.

do so would involve prioritizing functions that anger is thought to have played around 2 million years ago, when our ancestors were not highly cooperative, over the functions anger likely began to play around 12 thousand years ago, from when our ancestors are thought to have become intensely cooperative (Sterelny 2016). Therefore, even if the group level functions of anger evolved from the individual ones, we need further reason to take the phylogenetically eldest function of anger as 'basic' in our theorizing about human anger, given that we bear far more similarities to cooperative societies than to our solitary hominid ancestors. Theorists that take anger's function from the pre-cooperative stage of our evolutionary history naturally find anger of limited value, a vestige of our evolutionary past without clear contemporary benefit and ill-suited to present day society.

This type of story still doesn't establish which of anger's functions at the individual level is primary, if any, however. Is the epistemic prior to the behavioural function? or should they share the status of anger's primary functions? Similarly, although the story sketched above can make sense of the group level of functional analysis deriving from the individual level, a hierarchy is not easily established between the individual and the dyadic levels. Is the communicative function of anger characteristic of the dyadic level derivative of, or secondary to, the individual behavioural and epistemic functions anger plays?

Many take the communicative function of emotions to be secondary to the behavioural (Darwin 1872; Griffiths 1997). Behaviors came with characteristic facial expressions (showing teeth in preparation of attack for example). Although these expressions may no longer serve a behavioural function, the thought is that the expressions have been kept because they served a secondary communicative function. This explanation faces an immediate challenge, for there are many expressions that do not seem to be linked to behaviour. Smiling and frowning seem to be primarily communicative for example. The behaviour-first explanation cannot be the full story as it would seem that purely communicative features of emotions proved adaptive despite not reflecting any behavioural intentions. We are concerned with anger though, for which such a story might be more plausible: anger's involuntary expression might indeed be linked to behavioural tendencies. Even if this is the case, however, this fact alone is insufficient to commit us to anger's basic function being behavioural as opposed to communicative. To endorse this, one would need to hold a further commitment, that is, that anger's basic function is the most phylogenetically ancient function we can identify. We need to be given a reason for why anger should be

characterized along its presumed most ancient phylogenetic function, especially when this involves singling out functions that were selected for prior to two evolutionary shifts towards cooperation throughout our evolutionary history (Sterelny 2016).

Even if the group level functions of anger evolved out of functions described at the individual and dyadic levels, and if the communicative is, sometimes, phylogenetically younger than the behavioural, we might still have two distinct functions competing for the status of anger's 'primary' function: the epistemic, the behavioural. The typical way of construing anger's basic function is in behavioural terms, however. Anger picks up on cues so as to allow us to act on or in response to them. The epistemic serves the behavioural and the behavioural emerges as, plausibly, the most phylogenetically primary function of anger. These types of moves collapse communicative and epistemic functions into behavioural ones, or make them subservient to them. So long as one is responding to stimuli, however, anger is playing a basic epistemic role, and so long as one tends to act in certain ways in response to certain sets of stimuli, one's behaviour communicates something of importance to others. This suggests that attempting to pinpoint one adaptive function for anger at the individual level might be misguided, and that we should perhaps instead endorse the view that anger has multiple functions, neither of which is derivative of the other in our characterization of the emotion[37].

If we are prioritizing phylogenetic age, then at the individual and dyadic level, epistemic, behavioural, and communicative functions emerge as basic adaptations. If we do not prioritize phylogenetic age, which I have suggested we should not, given that we are interested in human anger, then the group level functions are also basic functions of anger, as they are adaptations to cooperative living situations. This means that the commitment to anger being biologically basic need not involve a commitment to anger having one unique adaptive function. This is plausibly the case, even if phylogenetic age is one's guiding priority, given that it is not entirely straightforward to collapse the communicative and epistemic into the behavioural.

Another concern, is that the simplistic construal of anger's basic behavioural functions, as related to attack, and by extension retribution, seems to ignore facts not only about human

---

[37] Cogley (2014) for example supports a tripartite conception of anger's primary functions, where these are: behavioural, epistemic, and communicative.

evolutionary history, as I have suggested, but to also ignore work on anger in other species. At its most primitive stage anger is thought to cue individuals to basic transgressions, to motivate attack and communicate the intention to attack. Such a view is supported by work on aggression in non-human primates, which are thought to be instances of anger homologues (Brosnan 2006). An initial caveat here is that the data might be biased towards an aggressive construal of anger given that this is its most salient behavioural manifestation. It is quite likely that anger's behavioural manifestations in animals are far subtler and more complex than basic emotion theory seems to imply. There is, for example, evidence that animals regulate anger behaviour, in particular towards conspecifics, such that interactions involve an elaborate process of mutual signaling aimed at adverting physical altercations (Bluhm 2006). This suggests that even in animals, the anger affect program is more complex than merely triggering attack behaviour, and that complex communicative functions of anger are present in many animals.

With humans, in any case, most psychologists grant that aggression is neither necessary nor sufficient to characterize anger (Averill 1983; Griffiths 1997). Attention to anger at subsequent stages of our evolutionary past tells a far more complex story. Anger doesn't just motivate aggression, in fact outright aggression against in-group members would presumably be particularly costly in cooperative living. Furthermore, as social and kin bonds became more relevant to communal life it is plausible that anger manifested in different, and less straightforwardly aggressive forms. As the range of stimuli that triggers anger shifted and expanded to include groups of individuals, as well as communal norms, such as inequity, our evolution into more cooperative living will also have led to an expansion in anger's behavioural profile. Similarly, anger's communicative function is likely to have shifted to mirror this. There is no reason to assume that the information communicated by anger in pre-collective stages of our evolutionary part, (presumably the communication of threat of attack) would remain the same in the drastically different adaptive environment of cooperative living.

I hope to have illustrated the difficulty in pinpointing exactly what is 'hardwired' in basic anger. I propose that we take anger to play three basic functions: epistemic, motivational and communicative. In so far as we are interested in human anger however, the epistemic, motivational, and communicative functions will have been adapted to collective living where anger has been observed to pick up on norm-violations, including inequity, and is best

thought of as playing a range of motivational and communicative functions beyond mere attack and threat.

2.3   Anger as Psychologically Basic

Under a psychological construal of 'basic', basic emotions are those which cannot be broken up into further emotion components. The thought is that the complex phenomena of human emotional life are made up of simpler building blocks. The building blocks are thought to be basic in part because they do not themselves admit of further decomposition. These basic emotions are also those which we would be more likely to attribute animals and human infants. These basic building blocks are the affect programs of typically five basic emotions – happiness, sadness, anger, fear and disgust. The idea, therefore, is not that anger cannot be analyzed into distinct components such as hedonic valence, physiological changes, facial expressions and appraisals, but that none of these components are psychological emotion categories. Emotions such as thrill, on the other hand, would be construed as non-basic because it can seemingly be analyzed as a blend of fear and happiness, both of which cannot be made sense of by separating them into further, more basic, emotion categories. Indeed, the idea that our complex emotional repertoire is produced by the blending and mixing of a few basic emotions is not new. The Stoics endorsed four primitive emotions: delight, distress, desire, and fear; Hobbes highlighted seven simple passions (Deonna and Teroni, 2012: 19), that arise from modifications of the simplest of passions, aversion and desire, while Descartes spoke of six, and Aquinas of eleven, basic passions (James 1997: 6). Both Hume and Locke spoke of some passions being simpler than others as well. (Schmitter 2016).

In psychology the idea has been popular as well. McDougall (1926) took basic emotions to fuse together to create more complex emotions such as admiration and awe. Tomkins (1963) and Izard (1977) similarly argued that various emotions combine with each other. Plutchik (1962) proposed a colour metaphor for the combining processes of emotions, where basic emotions can combine together to create an entirely new emotion with distinct phenomenal qualities from the emotions it originated from.

It is unclear how helpful such blending and combining metaphors are. Firstly, it is unclear how basic emotions are meant to combine to create secondary ones. Secondary emotions are typically seen as less modular, more dependent on cognitive input and promoting of more

diverse action profiles. It is unclear how basic emotions with reflex like motivational outputs combine to create a less modular emotion that has looser behavioural consequences. If they can do so, however, this raises the question of why non-human animals, who possess the relevant basic emotions, do not form secondary emotions that we typically take to be unique of humans. Another concern is how to make sense of the functions of secondary emotions, which are typically not thought to be survival oriented, given that their functions are meant to be inherited from survival oriented basic emotions. The blending or building block story seems to raise more questions than it answers.

If the relation between basic and secondary emotions is not one of blending, what is it? One option would be to draw the divide based on universal and basic biological function. On such an account, basic emotions are those that are universal, and have a basic biological function, while non-basic emotions are those that do not. Damasio (2005) for example has proposed that whereas basic emotions are hardwired and universal, secondary emotions are learnt during childhood, and more subject to cultural variation.

An issue with this type of story is that the domains along which the basic emotions are thought to be basic, involve value judgements as to which survival goals were of basic biological value. Basic biological value is typically construed entirely along individualistic lines, ignoring the crucial adaptive shift to cooperative living. This arbitrarily favours the classification of fear as basic, for example, while excluding love, or some precursor of it, from being granted this status.

Furthermore, we have seen that determining a hard-wired function for one of the most emblematic 'basic' emotions, anger, is much harder than might be thought. We have also seen that anger, and presumably other candidate basic emotions, might be best construed as universal only in a thin sense that anger phenomena across cultures and some animals share a phylogenetic ancestor. Our discussion so far has significantly weakened the extent to which basic emotions are hardwired and universal, making Damasio's strategy for distinguishing between basic and non-basic emotions unlikely to succeed.

Griffiths advocates that we draw the distinction between basic and non-basic emotions based on neural profiles, claiming that there is evidence for processing associated with basic emotions being localized in the ancient limbic system, while processing associated with

76

secondary emotions are thought to be located in more cognitive, higher, brain regions. Let's take a look at this proposal. Some affective neuroscientists have taken specific dedicated neural networks to underlie basic emotions (Panksepp 1988). The regions associated with emotions are typically the brain stem, the thalamus and hypothalamus, and the limbic structures, including the amygdala. These brain areas are construed as evolutionarily primitive and relatively encapsulated, resistant to learning, and associated with generating automatic behavioural responses. These regions are also directly connected to the autonomic and immune systems that monitor and adjust the body's homeostasis. Basic emotions are therefore relegated to the 'lower', ancestral, subcortical regions of the brain, while cognitive functions such as reasoning, learning, and planning are located in the phylogenetically young 'higher regions' of the cerebral cortex. This leads to a view that lends scientific support for the juxtaposition between emotion and cognition. The emotional brain is seen as primitive and hasty, distinct from the higher regions that support considered thinking, and rational evaluation. Victorian neurologist John Hughlings-Jackson (1879) writes that 'the higher nervous arrangements evolved out of the lower (to) keep down those lower, just as a government evolved out of a nation controls as well as directs that nation' (354).

The secondary, more social and cognitive, emotions are thought to rely on processing in the neocortex. According to the story Griffiths (1997) subscribes to, only the neocortex underwent significant changes throughout our evolutionary history, leaving the seat of affect programs, the limbic areas, unchanged and therefore functionally encapsulated. There is evidence, however, that the functional separation of limbic and cortical areas is a feature of our evolutionary past that we have long left behind. There is evidence, for example, that parts of the amygdala are more developed in monkeys than rats (Chareyron et al. 2011), suggesting that the subcortical regions evolved to become more complex. Pessoa (2014) takes increases in the size of the amygdala to reflect the degree of connectivity with other brain structures. The human amygdala is known to have extensive reciprocal connections to the 'higher' cortical regions, suggesting that the emotional brain, and the neocortex, may have evolved to become more interconnected. The fact that the higher and lower brain regions are 'reciprocally connected' is important to stress. Reciprocal connections are bidirectional, such that the direction of causal effect is two-way. Unlike the view Hughlings-Jackson conveys, it is not the higher brain regions that monitor and suppress the lower ones, but rather the functioning of the lower emotional brain regions depends on the neocortex and cognitive processing, and vice versa (Barrett 2006; Pessoa 2014)

This suggests that anger evolved to become less encapsulated. Instead of humans experiencing and displaying two different types of anger, one primary affect program anger, and the other secondary more cognitive anger, 'basic' anger seems to have been molded successively to adapt to changing, and ever more cooperative environments. This occurred as connections to the growing neocortex increased extensively, changing the nature of even the most 'basic' form of human anger to become more dependent on higher-level processing[38].

The distinction between basic and non-basic or secondary emotions seems to be harder to forge than initially thought[39]. Indeed, the category of 'basic emotion' should be rethought and qualified, instead of taken up as solid ground for philosophical thinking about the emotions. Collombeti (2014) for example recommends abandonment of basic emotion talk. Basic emotions are not the building blocks for secondary ones, except in the strictly phylogenetic sense that they are likely evolutionarily younger. As I have already stressed, it is unclear why phylogenetically younger emotions should be considered more basic in our philosophical accounts.

2.4   Taking Stock

I have critically evaluated the answer to the question of what anger is, that a basic emotion theory would provide. In doing so I have dissected the possible commitments of basic emotion theory, a theory that has been highly influential in the philosophical literature. Three main outcomes emerge.

---

[38] It is relevant to note that Griffiths was writing at the end of the 20th century, at a time when constructivist theories of emotion in neuroscience were yet to gain popularity. There is an emerging consensus that the human brain resists decomposition into emotional and cognitive parts. Recent work supports this trend by making it unlikely that there be a one to one mapping of psychological functions onto brain regions. Brain regions have been observed to demonstrate substantial pluripotency (the same region can underlie multiple functions) and degeneracy (different structures can perform the same function). Basic emotion theorists are therefore unlikely to find discrete neural bases for specific emotions (See Lindquist and Barrett 2012).

[39] One way to draw the distinction between basic and non-basic emotions, available on the story I have developed, is to say that basic emotions are those that have affect program ancestors, while non-basic emotions are those that do not. This would leave open the view that some emotions, nostalgia for example, might be entirely socially constructed, and not bear a relation to an affect program ancestor. I think it is unlikely that any emotion bear no relation whatsoever to some affect program(s) ancestor, although some emotions will presumably be more closely related to the relevant ancestor than others. A distinction could, then, be drawn between those emotions that are more closely related to affect program ancestors than others. This would make room for the proximity between human and non-human emotional phenomena to be captured, without committing us to empirically implausible distinctions between basic and non-basic emotions.

Human anger should not be equated to, or held to constitutively involve, the affect program of anger. This is because the universal anger affect program is best construed as an evolutionary ancestor, rather than a component of human anger. Human anger is, therefore, universal only in the thin sense that all human anger phenomena likely share a common evolutionary ancestor: an anger affect program that we most likely share with many non-human primate species. This explains why there are reliable manifestations of anger cross-culturally, as the capacity for some form of anger is hardwired. A heterogenous construal of anger's nature is in order, however, to account for the role played by exogenetic inheritance in human evolution. This, in turn, allows the plurality of anger phenomena observed across cultures to be accounted for, without restricting such plurality to differences along the timescale of upbringing alone.

We have seen that characterizing anger functionally should involve far more attention to the different adaptive environments early humans faced and not reduce anger to one 'original' or unique function. The communicative, epistemic and motivational roles of anger were highlighted as potentially irreducible, as were the different forms each of these might take once the evolutionary stress shifted towards cooperative living.

Once the concept of basic emotions is scrutinized, the distinction between basic and non-basic emotions, or basic and non-basic forms of specific emotions such as anger, becomes hard to sustain. I have argued that the basic/non-basic distinction is harder to substantiate than originally thought, perhaps especially in so far as we are concerned with a particular emotion. It seems misguided to think that humans sometimes respond with a pure form of ancestral 'basic' anger, that has remained relatively unaffected by the millions of years of shifting adaptive pressures and neural development. The thought that cognitive forms of anger can be separated clearly from basic forms of human anger seems mistaken.

2.5   A note on Natural Kinds and the use of Experimental Work

Griffiths (1997), has famously argued against emotions forming natural kinds. Anger, for Griffiths (1997) effectively breaks up into two different referents: basic and non-basic anger. According to Griffiths, our common use of the term 'anger' refers to the affect program of anger on the one hand, but also to more culturally inherited and higher-level forms of anger, on the other. Griffith's view is that the heterogeneity of things that fall under the folk term

'anger' do not constitute a natural kind, as they call for differential investigation: affective science may be the way of investigating the affect program of anger, while historical, anthropological, and sociological analysis, may be the adequate forms of investigating 'higher level' or 'cognitive' non-basic anger. What is true of 'basic' anger may not be true of more 'cognitive' anger and, therefore, the former concept is not theoretically useful, or projectable, in investigations of the latter type of anger, and vice versa. These commitments lead Griffiths to his famous position that folk psychological emotion terms do not refer to natural kinds[40].

In so far as my project is an interdisciplinary one, the success of Griffiths argument is relevant to my concerns, as it impacts whether or not I am warranted in relying on the affective sciences to inform my theorizing about human anger. If scientific work on anger targets a distinct phenomenon to that targeted by out folk concept of anger, as Griffiths claims, and our philosophical investigation starts from this folk concept of anger, then insights gained on the former will be ill-equipped to inform the latter, and vice versa. In other words, if the affective sciences investigate affect programs, and my philosophical project investigates the higher-level 'non-basic' anger, felt towards social injustices, then according to Griffiths, I might not be justified in taking empirical work on the former to inform my theorizing on the latter.

The nature of 'natural kinds' is contested, but Griffiths is concerned with whether emotion concepts form natural kinds of the sort amenable to *experimental* scientific investigation. His view is that our folk concept of anger only forms a natural kind of this sort in so far as the term refers to the affect program for anger. Higher-level forms of anger, to which the term also refers, are likely too diverse to form a natural kind amenable to scientific investigation according to Griffiths, although they might admit of investigation from the social sciences and humanities.

An initial line of response to Griffiths would be to accept that the folk concept of anger does not form a natural kind amenable to scientific investigation, but to hold that despite this, it forms a natural kind amenable to *philosophical* investigation[41]. If we follow Griffith's own

---

[40] Griffiths' argument against the natural kind status of the emotion, and particular emotions, has been the target of much critique. See Charland (2002), Roberts (2003: 14-36), Prinz (2006: 81-6), and Deonna and Teroni, (2012: 21-27).

[41] Charland (2002) and Deonna and Teroni, (2012: 21-27) both make a similar point regarding emotion types and the category of 'emotion'. Both 'anger' and 'emotion' form natural kinds in the sense that they capture sets of objects with related features, that are sufficiently well-defined for philosophical investigation (regardless of whether they are sufficiently well defined for scientific investigation).

liberal view of natural kinds, however, anger can legitimately be considered, for the purposes of philosophical investigation, a natural kind. On his liberal construal, natural kinds do not require necessary and sufficient conditions, nor do their constituent properties have to be fixed at all times (Griffiths, 1997). A natural kind is a category that 'brings together a set of objects with correlated properties' (188). The thought is that natural kinds capture reliable generalizations that function successfully in induction and explanation (Griffiths 2004). Counterfactually then, emotions like anger form a natural kind in so far as their absence as a relevant category would significantly hinder the success of our inferences and result in a loss of explanatory power.

Our folk concept of 'anger' refers to affective phenomena that feel a certain way, that arise in response to certain occurrences (harms and norm violations, for example), and which often trigger behaviour aimed at rectifying the situation that triggered it. This set of features is distinct from the set of features that characterize other emotions, such as fear for example. This suggests that it is an explanatorily useful category, and therefore a natural kind on Griffiths liberal construal. Imagine what a loss we would be at to explain the subjective state, and behaviour, of agents responding to harms, if we lacked the concept of 'anger'. We therefore have reason to think that anger refers to a natural kind amenable to philosophical investigation[42].

We might have reason to reject Griffiths' narrow claim, regarding emotions not forming natural kinds amenable to scientific investigation, as well. By arguing that no modular anger affect program survived our evolutionary history, I have blurred the lines between basic and non-basic types of anger. Therefore, the distinction between basic and non-basic anger that Griffiths makes use of to deny anger natural kind status, is unavailable. Without a clear distinction between basic and non-basic anger, Griffiths' claim that our folk concept of anger doesn't form a kind amenable to scientific investigation due to its split referents, loses some of its force.

---

[42] Griffiths seems to be aware of this initial type of response and writes that 'scientific understanding is not the sole or main goal of everyday life. Vernacular emotion concepts serve other purposes besides those of explanation and induction. The future development of these concepts is unlikely to be as simple as their refinement or replacement by the concepts best suited to scientific understanding' (Griffiths 1997, 228). In later work he even moves away from using the term 'natural kind', in an attempt to be more precise about his claims regarding the emotions: 'A more neutral substitute for 'natural kind' that carries many of the right connotations is "investigative kind" (Griffiths 2004: 907).

The story I have told of anger lends some support to viewing 'emotion' as a natural kind, as well. In the first instance, much as with anger, 'emotion' can be a category amenable to philosophical investigation even if Griffiths is right that it is not amenable to scientific investigation. Indeed 'emotion' has been taken to be a distinct kind of psychological state by philosophers (de Sousa 1987; Gordon 1987; Tappolet 2016), as well as by psychologists (Lazarus 1991; Frijda 2005; Scherer 2005). The cluster of features that is taken to characterize the emotion kind includes:

1. An evaluation
2. Valence
3. Physiological changes
4. Distinctive phenomenology
5. Connection to action tendencies

Philosophers add to these a sixth feature: intentionality, a notion which psychologists typically cash out in 'information carrying' terms. The fact that there is agreement on key features that characterize the emotion kind, across philosophy and psychology, suggests that 'emotion' satisfies the minimal conditions for kind status. That is, the category 'emotion' is explanatorily useful in capturing projectable features of a set of phenomena. Counterfactually then, much like with anger, if we lacked the concept of emotion we would be at an explanatory loss in our attempts to make sense of ourselves and others.

Charland (2002) writes that 'to the extent that emotions are a distinct kind of representational state, governed by their own special laws and regularities, it is plausible to view them as a distinct psychological kind' (522). The kind of representational state emotions are is typically cashed out in terms of emotions involving evaluative content for philosophers (Solomon 1976; Tappolet, 2000) or appraisals for psychologists (Lazarus 1991; Scherer 2005). The appraisal(s), or evaluative content, of an emotion individuates the specific emotion type it is. The distinctive phenomenology of emotion types is related to its evaluative content, or the type(s) of appraisal involved, which have a certain valence. Many think emotional phenomenology correlated to, if not partly constituted by, distinct physiological changes, including patterns of motor readiness or action tendencies. 'Emotion' therefore emerges as a category fit for philosophical investigation (Deonna and Teroni 2012a; Deonna et al. 2015).

Having blurred the distinction between basic and non-basic emotions in this chapter, however, Griffiths' claim that 'emotion' doesn't form a unified category fit specifically for

experimental investigation loses one of its main pillars. His thought was that the basic emotions deserve entirely different treatment to secondary or non-basic emotions, and that therefore the term 'emotion', as a term meant to capture both basic and non-basic emotions, is not a relevant natural kind fit for scientific exploration.

This chapter has cast significant doubt on the division between basic and non-basic emotions. I take myself to have cast sufficient doubt on Griffiths underlying empirical assumptions to warrant that we not exclude empirical work as prima facie irrelevant to our concerns, for fear of the empirical work investigating an entirely distinct phenomena to the one that is our philosophical concern. Throughout this project I take our common use of 'anger' to be a term sufficiently well-defined to take empirical work on anger to be prima facie continuous with, rather than orthogonal to, philosophical work on human anger. Relying on experimental work might be problematic for a number of reasons, including methodological ones, but their relevance shouldn't be ruled out a priori for the types of concerns Griffiths voices. At times I will also rely on experimental work that uses 'emotion' as a guiding category. I take myself to be prima facie justified in doing so, because 'emotion' forms a sufficiently coherent category for our philosophical investigation to be in principle informed by work in the affective sciences.

2.6   The Emotion of Anger

Philosophical theories of emotion differ on how to account for each of the cluster of features I mentioned above, and how to construe the relations between them. To provide my working definition of anger, I will need to situate myself amongst the main possible positions on the relations between the main features of emotions. On many specific issues I will remain unproblematically agnostic. Where I do not remain agnostic, my commitments will be largely uncontroversial. A brief survey of the main stances in the literature will make clear that the construal of anger I will propose involves the endorsement of uncontroversial commitments regarding the nature of emotions. This a beneficial feature of my account of the rationality of anger: it would be argumentatively weaker to offer an account that depended upon a controversial account of the target phenomenon.

I follow most contemporary philosophers of emotion in taking emotions to be intentional states. By intentionality, I mean that emotions can have intentional objects, that they can be

about things[43]. When you are angry at me, I am the object of your emotion. Individuals are the paradigmatic objects of emotion, but states of affair, (such as having had my wallet stolen), social objects (such as 'the government') and structural features (such as 'the patriarchy'), are also among the possible objects of emotions. These intentional objects are the emotion's 'particular objects'.

Particular objects are to be distinguished from the 'formal objects' of emotions, which I will turn to shortly. For now, I wish to highlight a distinction between two main types of particular objects: target objects and propositional objects (de Sousa 1987). Target objects are the entity at which your emotion is directed, me, for example, when you are angry at me. Propositional objects are facts or states of affair that emotions are directed at. When you are angry at me for having stolen your wallet, I am the target of your anger, and the fact that your wallet was stolen is the propositional object of your anger. Similarly, when you are angry at the government for cutting health care funding, the government is the target of your anger, and that health care funding has been cut, is the propositional object of your anger. Many emotions will have both types of particular objects, while others might only have one. When you are angry at states of affair, that sexism is a prevalent structural feature of society, for example, your anger has a propositional object, but need not have a target. On the other hand, you might be angry at the government, or your mother, without having a propositional object for your emotion.

How emotions represent their particular objects is a divisive point in the philosophy of emotion. Feeling theories, such as William James', equate emotions to physiological changes. Such views have been charged with denying emotions intentionality (Deonna and Teroni, 2012; Tappolet 2016)[44]. Others have however argued for neo-Jamesian views that seek to make feeling theories and emotional intentionality compatible (Prinz 2006)[45].

Strong cognitivists take emotions to represent their objects conceptually, much like beliefs do (Solomon 1976; Nussbaum 2016). Cognitivist or Judgement theories of emotion have been criticized for being too cognitively demanding, and by consequence excluding human infants from having emotions. They have also been criticized for neglecting the felt aspect

---

[43] I remain agnostic on whether there are some cases of objectless emotions, but note that often, the intentionality of emotions is taken to be a definitional feature which distinguishes the kind 'emotion' from other affective kinds that lack objects, such as feelings, for example pain, and moods, such as gloominess (Deonna and Teroni 2012: 4).

[44] See Ratcliffe (2005) for a push-back against this common charge against feeling theories.

[45] See Goldie (2006) for a critique of Prinz' view.

of emotional experience[46]. The main insight of cognitivism over feeling theories remains, however, as most contemporary views endorse some form of weak cognitivism to account for emotional intentionality. Some take emotions to be belief-like states that fall short of full-blown beliefs (Greenspan 1988), others characterize them as evaluative construals (Roberts 2003), or perceptions of value (Tappolet 2000; Döring 2009). In line with this consensus, my construal of anger will involve a form of weak cognitivism to account for the emotion's intentionality.

Emotional intentionality is characterized as Janus-faced or double. Emotions not only have particular objects of the sort outlined above, but also have a formal object, typically an evaluative concept or property that emotions of the same type share (Kenny 1963; Mulligan 2007; Teroni 2007). Danger in the case of fear, for example, and offence or wrong in the case of anger. Any X that I can have emotion E about, is a particular object of E, whereas the formal object of E is the property which I implicitly ascribe to X by virtue of having E about X (Kenny 1963). How to construe the formal object is a subject of debate. The formal object of emotions supplies, at least in part, the evaluative component of an emotion. The formal object also plays a role in establishing the correctness conditions for emotions, as an emotion will be appropriate, or justified, when its formal object is one that is in some way instantiated by the particular object of the emotion. This has led most philosophers working on the emotions to endorse some form of realism about value, a trend which I follow (Deonna and Teroni 2012; Tappolet 2016)[47].

How exactly the evaluative property is represented in emotional content is debated. Perceptual theorists, and other cognitivists, tend to think that emotions have evaluative representational content (Döring 2009; Tappolet 2016), while Attitudinal Theories take

---

[46] See chapter one of Roberts (2003) or chapter five of Deonna and Teroni (2012) for thorough critiques of Judgement or Cognitive theories. For more nuanced versions of Judgement theory, that aim to keep emotional phenomenology center stage, see Solomon (2003) and Nussbaum (2004).

[47] Following perceptual theorists, as well as most contemporary philosophers of emotion, I will assume some form of value realism (Deonna and Teroni 2012). Generally, these can take either response-dependent or response-independent forms. Fitting-attitude analyses, where objects have the evaluative properties they have by virtue of a particular emotional response being fitting, are the most popular type of the first form of realism. Response-independent forms of value realism typically employ a notion of supervenience whereby evaluative properties are thought to supervene on non-evaluative natural properties (Deonna and Teroni 2012). The latter type of realism avoids issues of circularity that have plagued fitting-attitude analyses, but faces the challenge of providing a detailed account of the natural properties in question, as well as the nature of the supervenience relation. I leave metaethical issues aside, and move forward with a commitment to some form of realism about evaluative properties, while remaining agnostic about how this is best cashed out. Importantly, I lean towards an account where evaluative properties supervene on non-evaluative properties, whether the evaluative properties are themselves response-independent or not. See Deonna and Teroni (2012: Ch.4) for a survey on the relation between emotion and value.

emotions to be evaluative attitudes where the evaluative component of emotions is given by the type of state the emotion is, rather than by its representational content (Deonna and Teroni 2012). The thought here is that an intentional state's mode, or attitude, comes apart from its content. Truth, for example, is a feature of the attitude or mode of belief, as truth need not feature in the content of beliefs for one to believe a certain proposition true. Similarly, one need not represent a lion as dangerous, but merely represent a lion under the mode of evaluating it as dangerous – that is, fearing it – for Attitudinal Theories (Deonna and Teroni 2012).

I remain agnostic between whether evaluative properties feature in the representational content of anger, or whether evaluative properties are actually a feature of anger's mode of presentation, or attitude. However, I am committed to anger having evaluative *phenomenal* content. That is, anger's phenomenology is evaluative in that it feels as though the particular object of anger instantiates the relevant formal object.

Emotion types have a distinctive phenomenology that is often taken to be partly constitutive of making that emotion the type of emotion it is. The way anger feels is distinct from the way fear or joy feels, and indeed the phenomenological salience of emotions is perhaps their most distinctive feature. Emotion types are also thought to be associated with certain physiological, as well as action tendency, profiles. Motivational theorists (Frijda 2005; Deonna and Teroni 2015) take the physical states of action readiness, that characterize each emotion type, to give emotion types their distinctive phenomenology. On such accounts anger feels the way it does because of the motivational state one is put in when experiencing the emotion[48].

Motivational views have been criticized on the basis that many emotions have a wide range of related action tendencies, and some such as happiness don't typically have any at all. Indeed, these views have been critiqued for endorsing a view whereby emotions are characterized by motivational modularity (Tappolet 2009). My critique of basic emotion theory, in this chapter, should have left us skeptical of any strict modularity in the realm of human emotions, be this motivational, or otherwise. Motivational modularity sits well with a view of anger as playing a particular hardwired behavioural function. This is an assumption I have critiqued in this chapter. I take anger to be causally, rather than constitutively, related

---

[48] Note that attitudinal theorists need not be motivational theorists, they can take emotional attitudes to be characterized by distinctive physiological changes, or phenomenal states, separate from motivational ones.

to action tendencies. We will see in chapter 3 that anger-caused actions are hard to fit under one goal or aim.

I do not intend to defend a particular theory of emotion or propose a new one. It will be a strength of my account of the rationality of anger that it not hinge on a specific theory of emotions. That being said, there are a few uncontroversial commitments with which I am aligned, some of which have already been highlighted.

Firstly, we have seen that emotions can have intentional objects, they are about things and states of affair in the world. The instances of anger that we will be concerned with have such objects. I will take anger to therefore have cognitive content in this minimal sense. It is plausible that anger's representational content can be conceptual, and non-conceptual, depending on the type of object it is felt towards (see Salmela 2014). I take anger, much like most emotions, to be an embodied state that feels a certain way in part due to physiological changes. I have said that emotions have evaluative *phenomenal* content, that is, they feel as though their particular objects instantiate the associated evaluative properties or formal object. Anger, therefore, involves a feeling of having been harmed, and involves apprehending its target as in some way responsible for this harm. The phenomenology of anger mostly has a negative valence, it is typically unpleasant to be angry, and the formal object of anger is a negative evaluative property: offence, harm or slight. However, anger also involves feeling that the triggering situation can be in some way changed, and that one does not deserve to be treated in such a way. Anger's valence, therefore, is unlikely to be entirely negative[49].

Motivational theorists would be keen to characterize this aspect of anger's phenomenology in terms of the motivational states that they take anger to involve. From what we have seen about anger so far however, I am skeptical of any account of emotional nature and phenomenology that ties it too strictly to specific motivational states. Although not all emotions have strong ties to conative states, anger is thought to involve a conative

---

[49] This is in line with the first-person accounts of individuals living under oppressive conditions, that we saw in chapter 1. It is also in line with work in appraisal theory of emotion. According to appraisal theory, emotions constitute, or are caused by, a number of appraisals along different dimensions. On such theories, anger typically involves the following appraisals: that one of the subject's aims or goals has been blocked, that an offence or injury has occurred, that someone is to blame for the cause of anger, and importantly, that one has the ability or power to cope with the triggering event, and even change it (Scherer 2005; Lerner and Tiedens 2006).

component, and therefore I must say more about it. My next chapter will address what the nature of this conative component might be.

2.7 Conclusion

I have argued that anger should only be construed as a basic emotion in a thin sense, namely that it has a primitive anger affect-program as an evolutionary ancestor. Attention to human evolutionary history revealed anger to best construed as a heterogenous kind. Not only has human anger evolved, in the classical sense, to enable the communal living of our highly cooperative species, but anger has also evolved exogenetically, such that cultural and social practices partly constitute anger's heterogenous nature.

I argued that this account of anger significantly undermines the main attempt to deny the category of 'anger', as well as the category of 'emotion', natural kind status. In light of this, I have argued that I am prima facie justified in taking the subject matter of empirical and philosophical work on the emotions to be continuous, and therefore I am prima facie entitled to supplement my philosophical project with work from the affective sciences. I ended by outlining my uncontroversial initial commitments regarding the nature of anger in relation to the main positions within contemporary philosophy of emotion.

In this chapter I have challenged the case for thinking that any form of human anger is a primitive reflex-like state aimed at attacking or harming offenders. In so doing I have laid the foundation for skepticism regarding anger's presumed constitutive tie to a desire for retribution. Anger has emerged as a state capable of playing a variety of functions. In the next chapter, I will provide further argument against the construal of anger as constitutively tied to a desire for retribution. We are particularly concerned with anger's ability to enact social change, as this is the positive role that has been most forcefully advocated of anger. I will, therefore, focus on anger's conative component in scenarios of this sort.

3

Anger's Desires: Recognition and Retribution

'It is no accident, then, that the emotion which accompanies
the first steps toward liberation is, for most women, anger....Through the
exercise of your anger, as you see its efficacy and thus your own,
you gain strength.'
-Susi Kaplow in, 'Getting Angry

What of the conative component of anger? We saw in chapter 1 that construing anger to involve a desire for revenge, payback or retribution, has a long philosophical history, indeed Nussbaum calls this the 'traditional' view of anger (2015: 41). We have seen that Aristotle characterizes anger as 'a desire accompanied by pain for conspicuous revenge caused by a perceived slight' (*Rhet.* 1378a31-33). Both Stoic and Buddhist philosophers strongly condemn anger, as they take it to involve a desire to lash out and harm the object of anger (see Bommarito 2017). We also saw that the punitive construal of anger is alive and well in modern scholarship as well. Ben-Ze'ev (2000) for example writes that 'the urge to attack is essential to anger, even if it is expressed in a nonstandard aggressive act' (384). Similarly, Pettigrove (2012) takes anger to 'involve the desire to lash out at its object or to see that object hurt' (358).

In chapter 2 I argued against an extremely influential theory in the affective sciences, basic emotion theory. Basic emotion theory could be seen to lend support the traditional construal of anger. The thought was that if anger is best construed as a modular affect program, with set behavioural tendencies hardwired from our evolutionary past, anger may indeed be a state that aims for attack and retribution. As a modular affect program, anger's behavioural tendencies would be mostly automatic, and typically insensitive to conscious control or cognitive input. This would lend support to a construal of anger along the lines of those who condemn it, as anger would trigger action largely independently to the agent's wide concerns or aims. My argument in chapter 2 developed a systematic and empirically informed critique of this take on anger. I argued that anger is not a modular affect program, and that it is best construed as a heterogeneously constructed kind that admits of multiple behavioural functions.

That being said, it might be the case that retribution is still the main behavioural manifestation of anger in situations of social injustice. Given that I have given some evidence for the behavioural pluripotency of anger, it is important to hone in on the particular cases of anger we are concerned with. In this chapter, therefore, I focus on anger under conditions of social injustice, broadly construed. Both Nussbaum (2015, 2016) and Pettigrove (2012) have launched recent attacks on the efficacy of anger advocated by feminist philosophers (Frye 1983; Narayan 1988; Jaggar 1989) and civil rights activists (Adams 1995) in struggles for revolutionary justice. These attacks hinge on their endorsement of the traditional construal of anger being constitutively tied to a desire for retribution. For now, by 'desire' I will mean a disposition for particular actions, in this case dispositions to act in ways that enact retribution[50]. In this chapter, I will focus on Nussbaum's argument in particular, and argue that it is unsuccessful in establishing that anger is ineffective in fighting against social injustice. In doing so, the traditional conception of anger as involving a constitutive desire for retribution will come under scrutiny.

Nussbaum takes anger to be typically ineffective, and often counterproductive, to the aims of the oppressed, in large part due to this conceptual link between anger and retribution. In

---

[50] This is the canonical and conservative construal of desire (see Schroeder 2017). Neither Nussbaum (2016), nor Pettigrove (2012), explicitly specify what construal of desire they have in mind. We, therefore, have prima facie reason to assume they are working with the canonical construal of desire, especially given that their condemnation of anger centers around the actions they take anger to cause. It is clear that they have in mind a construal of desire that grants anger a close link to the behaviors that it causes.

section 2 I detail Nussbaum's view and outline these two commitments of hers – the *nature claim*, and the *inefficacy claim*. We will see that the *inefficacy claim* is held for two reasons that fall out of the nature claim, both of which involve questionable empirical assumptions.

I will survey empirical work that challenges both reasons for holding the *inefficacy claim*, in so doing we will uncover reasons to question the *nature claim*. I argue that the empirical work supports moving away from a traditional construal of anger as conceptually tied to retribution, and that it instead fits a view that takes anger's desire for recognition seriously.

This chapter will contribute to the meeting of my first and second desideratum. Investigating anger's conative component(s) will contribute both to the provision of an account of anger's nature, as well as to the provision of an account of anger's instrumental rationality. This chapter lends support to the instrumental rationality of anger in fighting against social injustice. We will see that this practical utility is often linked to a desire, in anger, for recognition, rather than retribution.

3.1   Retribution

Nussbaum (2015:46) makes explicit that she takes the punitive aim of anger to be constitutive of the emotion, rather than being merely causally related to it. The desire for the perpetrator's suffering becomes a conceptual part of anger, rather than a feature of the stereotypical expression to which it is causally related. The challenge I build against Nussbaum will give us cause to doubt not only the strong constitutive link, between anger and retribution, but the robustness of the supposed reliable causal link between the two as well.

What exactly is anger's retributive aim? Nussbaum's thought is not that anger conceptually involves a desire for violent revenge or to personally harm the offender, rather, 'anger involves, conceptually, a wish for things to go badly, somehow, for the offender in a way that is envisaged, somehow, however vaguely, as a payback for the offense' (Nussbaum 2015: 46). For example, in anger at a personal betrayal, one may wish for the traitor's life to go badly, yet not desire to have nothing to do with making this the case[51]. Nussbaum's view of anger involves the following nature claim:

---

[51] This might suggest a departure from the canonical, action-based, dispositionalist reading of desire that I have assumed Nussbaum is committed to. According to this quote, the desire for retribution seems not to involve

Nature Claim: Anger constitutively involves a desire for payback or retribution.

Nussbaum argues for two related points in her attack on anger. Firstly, that given the nature claim, anger is either intrinsically irrational or immoral, and secondly, that contrary to much feminist and activist thinking, anger is ineffective at fighting oppression. Her reasons for the latter are relevant to our concerns, for if her argument succeeds then we have reason to seriously doubt the instrumental rationality of anger in the contexts in which its utility has been most forcefully highlighted, namely in fighting against social injustice.

For Nussbaum, anger is, by its very nature, either irrational or immoral. It is irrational in the sense that inflicting pain upon a perpetrator when in anger will not literally 'undo' the wrong one has suffered. An agent who believes the contrary is guilty of 'magical thinking' and irrationality (2015: 47-48). The only way to avoid irrationality is to construe the payback as restoring one's status after it has been injured, by having incurred a slight. This is in line with the Status or Honour reading of Aristotle on anger. Anger on this view is concerned with status-ranking and 'a retaliatory strike back is thought to restore the balance of status' (Nussbaum 2015: 48). The problem with status-focused anger, for Nussbaum, is that it is immoral because it involves a 'narcissistic error' (2015:51), an obsessive focus on one's standing relative to others (2015: 45). There is however one domain where Nussbaum (2015) grants that slights do lower one's status, and that to be preoccupied with such status injuries is not a moral vice:

> Discrimination, for example, on grounds of race or gender, is often conceived as an injury that really does consist in down-ranking, and there is truth to this, just in this special sense: discrimination involves a denial of a special status of equal dignity, and this status has intrinsic value. (50)

---

dispositions for retributive actions, but a disposition to derive pleasure from, or be satisfied in, things going badly for the offender. This could also be cashed out in altogether non-dispositionalist terms, where the desire for retribution involves seeing (perceived) retribution, or payback, as good. On both these pleasure-based or good-based theories of anger's desire, the disposition to enact retribution oneself is not central. A first point to note is that Nussbaum's quote is best taken as evidence of endorsement of a *holistic* view of desire, one that incorporates pleasure or good-based considerations, rather than one that excludes dispositions for action. That being said, I take my critique of the traditional view of anger (as involving a desire for retribution), to be a strong one, whichever view of desire is taken, so long as the desire has some connection to action. On an exclusively pleasure-based, or exclusively good-based, account of desire, anger would not involve *dispositions* to act retributively, but may well still involve the privileging, or recommendation, of actions aimed at retribution. Desires are things agents often act on, or in light of, whichever theory of desire is favoured. Traditional views of anger that endorse a non-action-based theory of desire would, therefore, still have anger be a state on which retributive actions are often based. This chapter challenges the constitutive link between anger and any such desire for retribution.

But, she goes on to say, 'the idea that denials of equal dignity can be rectified by bringing the injurer low is a false lure' (2015: 51), this is because 'reversing positions through payback does not create equality. It just substitutes one inequality for another', according to Nussbaum (2015: 51). Much more could be said here. Surely lowering the rank of those in power need not involve the reversing of positions with those oppressed, but merely equalizing them. It is unclear why such an aim would make payback morally problematic. Pursuing such lines of argument will not be my concern, however. I will be concerned with building a challenge to the view that anger constitutively involves such desires for payback or retribution, in the first place.

Nussbaum's commitment to the *nature claim* leads her to hold that anger is ineffective at fighting against social injustice. On the inefficacy of anger, Nussbaum argues that 'non-anger and a generous disposition are far more useful' to revolutionary justice (2016: 228). In a similar vein, Pettrigrove (2012) advocates meekness over anger in the fight against social injustice. Nussbaum doesn't just make comparative claims regarding anger's inefficacy however, she also argues, largely through counterfactual forms of argument, that leaders of successful social movements, such as King and Mandela, were effective precisely because they did not act on their anger (2016: Ch. 7)[52]. These comparative, as well as counterfactual, moves make it clear that, in Nussbaum's view, there is something particularly ineffective about anger in the fight for social change. This claim can be summarized as follows:

Inefficacy Claim: Anger is typically ineffective at fighting social injustice.

This claim is by no means limited to Nussbaum (2015) and Pettigrove (2012), and their Stoic ancestors. Given that the construal of anger along punitive lines is widespread, the *inefficacy claim* is plausibly implicitly endorsed by most philosophers who recommend against anger in general[53]. The inefficacy claim is also common in everyday life; we are often advised to

---

[52] This reading of these historical figures is open to criticism, Adams (2007) and Cogley (2014) for example read these political figures as evidence of exactly the contrary: anger being effective in promoting social change.
[53] As I suggested in chapter 1, there is room to argue that retributive anger is actually extremely effective at bringing about social change, and I say more on this throughout the chapter, but, this is not my focus. My focus is on challenging the traditional construal of anger as retributive, on which most of anger's bad reputation hinges.

repress and avoid anger on account of it 'doing no good' to our cause, and for fear of it being counterproductive to one's aims[54].

What are the specific reasons for holding the *inefficacy claim,* however? The answer becomes clear when we pay closer attention to the *nature claim.* Firstly, as anger is constitutively aimed at payback, retributive rather than conciliatory actions are predicted of those in anger. If anger involves a constitutive desire for payback then the disposition for retributive actions is constitutive of anger. Retributive actions are both morally problematic for Nussbaum, and risk being counterproductive (Nussbaum 2016: 1).

When enraged African-Americans flooded the streets of Chicago following Martin Luther King Jr.'s assassination in 1968, for example, 125 fires were set, 210 buildings were damaged and countless stores were looted. In response to the riot, over 10,000 police and 5000 soldiers were brought in. Many African-Americans were killed, injured or incarcerated, the city of Chicago suffered a food shortage, and areas destroyed by rioters were knocked down and many remain to this day undeveloped. Racial segregation intensified in the riot's aftermath (Risen 2009). Besides committing harms against Chicagoans in general, these actions arguably left the African-American community worse off in the short run, as it provoked backlashes that intensified oppression. This lends support to Nussbaum's view that anger simply doesn't further your goals as it is 'incompatible with forward-looking pragmatism' (2016: 230–233).

There is, however, clear room to argue for the historical efficacy, and even justifiability, of destructive and aggressive actions in promoting social change[55]. Indeed, the Chicago riots are thought to have played a crucial role in leading to important longer-term victories for the civil rights movement (Risen 2009). The efficacy of violent revolutionary tactics is, however, not the focus of this chapter. I will be challenging the assumptions that underpin any view that takes anger's efficacy to depend solely on its link to payback, as well as any view that sees anger's efficacy as dependent on it becoming morally problematic.

---

[54] See Srinivasan (2018; 124-125) for further examples.
[55] Recall the famous example of the Haitian 1791 slave revolt; 100,000 slaves revolted against their masters and raped, tortured, pillaged, mutilated and killed most of the white population (around 4000 people), gaining control of one third of the island. The revolt was crucial to the emancipation of Haitians, by 1794 slavery had been formally abolished and by 1804 Haiti was an independent country (Knight, 2000).

Independently of the specific actions taken, commitment to the *nature claim* gives us reason to think that mere displays, or communications, of anger risk setting back goals for social change. This is because anger presumably communicates dispositions to engage in retributive actions. This brings me to the second reason anger is taken to be ineffective in revolutionary justice: that communicating anger, for those committed to the *nature claim*, involves communicating the desire for retribution, and that this is likely to inspire animosity in those that are the targets of anger. Nussbaum (2016) says that anger 'breed(s) mistrust' (233) and 'increases the anxiety and self-defensiveness' in its targets (230). Similarly, Pettigrove (2012) writes that anger communication is typically counterproductive due to 'triggering a defensive response' in its targets, that prevents them from appreciating the causes of anger (367). In a best-case scenario, the communication of anger, in so far as it communicates a desire for even mild retribution, is unlikely to breed openness to cooperation in its targets. In a worst-case scenario anger risks perpetuating an 'endless cycle of blood vengeance' by triggering retaliation (Nussbaum 2016: 1).

There might be other reasons for endorsing the *inefficacy claim*, but I will focus on the two I have outlined above. The first is a reason that pertains to the actions of those in anger, while the second is a reason that pertains to the responses of the targets of anger. They can be summarized as follows:

> In-group Reason*:* Anger motivates retributive actions on the part of those in anger.
>
> Out-group Reason*:* Anger antagonizes those at whom anger is directed.

Each reason is in effect an empirical claim that can be tested against experimental work. Before I turn to do precisely this, I would like to note the one exception Nussbaum makes, one case where she takes anger to be an effective (and moral) way of promoting social change, cases of what she calls 'transition-anger'. Nussbaum characterizes transitional anger as anger that is not retributive, and that focuses on 'brotherhood', 'justice', 'reconciliation and shared effort' instead, typically motivating constructive actions (2015: 53-54). This is the type of anger Nussbaum takes King to experience and express in his speeches (54). Nussbaum isn't clear on whether transitional anger is a distinct species of anger on her view 'is Transition-Anger a species of anger? I really don't care how we answer this question' (54). What is clear is that she takes it to be a 'borderline case' that is 'rare and exceptional', and only present in individuals with superior 'self-discipline' (54).

In sum, Nussbaum's view on anger takes it to be typically ineffective (as well as typically immoral) in struggles for social justice. When anger is effective it is morally problematic, by harming, or threatening to harm others. The only room made for the permissible efficacy of an emotion akin to anger, is the special case of transition-anger, which is exceedingly rare and hard to cultivate.

3.2   The Efficacy of Anger

If the question on the table is 'should the oppressed avoid anger?' We have seen that Nussbaum, along with other like-minded thinkers, would respond strongly in the affirmative. Feminist thinkers on the other hand have issued powerful responses to the above question in the negative. The oppressed should not avoid anger, on their view, for a number of reasons. These include anger's psychological, epistemic, as well as practical, utility in resisting oppression. Here I focus on the practical. Feminist thinkers are likely opposed to the *inefficacy claim* in the realm of social justice, as they take anger to be crucial in motivating politically beneficial action. We saw that Lorde (1997: 280), for example, wrote, that 'every woman has a well-stocked arsenal of anger potentially useful against those oppressions, personal and institutional, which brought that anger into being'.

Those who have argued for the political value of anger, have rarely engaged in a head on denial of the *inefficacy claim* however. This is likely in part because such a task is seen as largely an empirical one (Lepoutre 2018: 3; Srinivasan 2018: 127). The explicit denials of the inefficacy claim that can be found are more akin to calls to arms than developed accounts of anger's efficacy[56]. Recall Lorde's claim that 'focused with precision it (anger) can become a powerful source of energy serving progress and change' (1997: 280). Little is said by way of an account for how anger is to play such roles. Furthermore, Nussbaum grants that anger can cue us to injustices, and act as an initial motivator or spur to action (2015: 55), but she denies that acting on or in anger beyond this is effective to promoting change. This is the *inefficacy claim* I will challenge.

---

[56] Lepoutre (2018) is an exception to this gap in the literature. His argument relies upon close readings of public speeches of activists however, whereas mine relies on experimental work in psychology.

My challenge to the *inefficacy claim* will call for a rethinking of the conative component of anger, as I will argue that the empirical work supports anger often involving an aim for recognition, as opposed to retribution. By aiming for recognition, anger will involve an aim for an epistemic change in its target. Although the role of anger as an epistemic resource in conditions of oppression has received a lot of attention, the focus has been on anger being a crucial mode of perceiving injustices (See Narayan 1988, and Jagger 1989, for example). This focus doesn't place much emphasis on anger as a communicative phenomenon, that has effects on its targets, or on third party observers. In so far as the communicative dimension of anger has been addressed in the feminist literature, the focus has been on the widespread dismissal or lack of 'uptake' the emotions of oppressed groups are met with, and what we can learn from patterns of such dismissal (Frye 1989; Campbell 1994). For these thinkers then, anger should not be avoided by the oppressed because doing so would deprive them of a crucial epistemic resource for picking up on and mapping oppression. A more practical constructive role for the communication of anger remains underexplored[57].

The idea that anger aims at acknowledgement, or recognition, of an injustice by those that perpetrated them, is prevalent in the philosophical literature on Truth and Reconciliation processes (Walker 2006; Chakravarti 2014). This literature deals primarily with the role of anger in restoring relations *after* injustices, however, rather than anger's efficacy in combating ongoing ones. Recall that moral philosophers that give emotions a central role in moral action have highlighted the epistemic rather than retributive aim of anger. Adam Smith wrote that anger's aim is not to cause its target to 'feel pain in his turn... as to make him sensible, that the person whom he injured did not deserve to be treated in that manner' (1976: 95-6). In a similar vein, we saw that Darwall (2013) takes angry attitudes to 'not seek to diminish, humiliate, insult, or damage their objects. The acknowledgment they seek is of a reciprocal standing' (Darwall 2013: 89).

This chapter will lend support to recognition being one of anger's main aims. I will argue that anger's efficacy is linked to this aim for recognition, and that recent empirical work supports this claim. This will, in turn, recommend the rejection of the traditional conception of anger as constitutively tied to retribution. I, therefore, provide a head on challenge to the

---

[57] Exceptions in the philosophical literature include Srinivasan (2018) Cogley (2014), and Lepoutre (2018). All of whom have mentioned the communicative function and epistemic aim of anger.

inefficacy claim, that is currently lacking in the literature. I begin by building an empirical challenge to each of the reasons for holding the *inefficacy claim*.

3.2.1 The In-group Reason

Anger is ubiquitously seen as a crucial motivator for political action (Spring et al. 2018). The field of collective action research takes there to be two main pathways by which collective action is motivated: the anger pathway, and the instrumental reasoning pathway (van Zomeren et al. 2012; Wlodarczyk et al. 2017). The anger pathway involves the experience of group anger being triggered by a situation of unfair in-group disadvantage, while the instrumental reasoning pathway involves reasoning about how effective one's group will be at changing the unjust situation. The anger pathway is driven by appraising situations as unfair, while the instrumental reasoning pathway is driven by evaluating the amount of social support for action one expects. This involves reasoning about how successful the group is likely to be at ensuring change through collective action (van Zomeren et al. 2012).

Given the In-group Reason for holding the *inefficacy claim*, a supporter of this claim would plausibly expect the actions typical of each of these pathways to differ. Nussbaum would likely expect the instrumental reasoning pathway of collective action alone to motivate effective, and morally unproblematic, actions in the pursuit of justice. Retributive collective actions that are potentially counterproductive to this aim would, on the other hand, be expected of the anger pathway. Experimental work on collective action seriously challenges these predictions however.

In a key study, German university students were surveyed regarding a real-life situation where the state had mandated an increase in tuition fees. Students were asked to indicate how likely they would be to participate in different actions against the tuition rise. The action options were grouped into three types; a) 'constructive actions' such as flyer dissemination, petition signing and demonstrations, b) 'destructive actions' such as arson attacks on university buildings or private property and c) 'intermediate type' actions that disturb events where tuition-rise advocates appear, such as blocking university buildings or public roads (Tausch et al. 2011). Adherents to the *inefficacy claim* would predict the anger pathway to mainly motivate actions of type b), destructive actions, while predicting instrumental reasoning to

be the main pathway for motivating actions of type a), constructive actions. On the contrary, however, both anger and instrumental reasoning were found to be positively related to engaging in type a) actions, i.e. constructive actions[58]. Crucially, anger was not found to motivate actions that involved enacting payback in any straightforward sense, as anger motivated actions to change the tuition fee policy, rather than harm those who implemented it.

Similar results have been observed in a range of experimental set ups (van Zomeren et al. 2004; Tausch et al. 2011). For example, similar results were found in studies involving Muslim Indian minority communities in conflict with the Hindu majority. The Muslim community is one of the most disadvantaged communities in the country in terms of education, income, employment and political representation (Basant 2007). The self-reported levels of anger amongst the Muslim community in the riot-prone city of Aligarh, were found to be unrelated to any support for violent actions against the dominant majority. This suggests that the results of the study conducted on students plausibly extend to real world situations of historical conflict. Other studies have found anger in situations of group conflict, such as Isreal-Palestine, to promote both destructive actions, as well as constructive actions, against the out-group (Halperin 2008). The evidence above speaks against our first reason for holding the *inefficacy claim*, as anger is not shown to motivate destructive actions over constructive ones.

Some authors, Pettrigrove (2012) for example, have cited studies that show anger to motivate punitive and aggressive actions. Nussbaum and supporters of her view would likely try to explain away the constructive effects of anger that I just illustrated as exceptions to this norm. A few points on this. Firstly, the experimental evidence relied upon to establish the *inefficacy claim* are almost exclusively from studies done on individuals or in interpersonal settings (Pettigrove 2012: 362; Spring et al. 2018), therefore we should be skeptical of how they translate to the inter-group dynamics we are concerned with. Furthermore, in section 3.3.1 we will see that there is ample evidence of anger being constructive in interpersonal settings as well, which theorists arguing against anger's efficacy have neglected. Most importantly however, I have not denied that anger can have negative effects, the point is that constructive

---

[58] Anger and instrumental reasoning were measured by asking participants to rate how strongly they agreed with statements like: 'I am furious about tuition rises', 'I am irritated by tuition rises' in the case of anger, and: 'I think that students can stop the introduction of tuition fees', 'I think students have already lost the fight against tuition fees', for instrumental reasoning about group efficacy.

effects might be far more common than those who condemn anger grant, and indeed constructive effects might be just as typical, or even more so, of anger than negative ones[59].

How do we account for the variance in anger's motivational tendencies however? Experimental work suggests that key factors *moderate* the effects of anger. In psychology, moderators are crucial to determining when certain effects hold. Moderators are typically contextual variables that influence which effects are observed. Contextual moderators are likely crucial to determining whether anger will motivate constructive or destructive actions.

The Tausch et al. (2011) study for example, found destructive type b) actions to be favored when the group had low confidence in their ability to change their predicament. This suggests that taking a situation to be unchangeable may be a key factor in motivating destructive behavior. Indeed, there is wide ranging evidence that punitive actions are favoured in situations where change is unlikely. This might be because the out-group is unresponsive to attempts to change the situation (Bandura 2000). Indeed, some have called this the 'nothing to lose' phenomenon (Scheepers et al. 2006) as the low status group has little to lose in responding to injustice aggressively, seeing as their situation is unlikely to change by any other means. The perceived changeability of the out-group, therefore, seems to be a key moderator of anger behaviour. This suggests that whether the In-group Reason holds or not is heavily dependent on how changeable the outgroup is perceived to be[60].

In sum, we have seen the In-group Reason for holding the Inefficacy Claim to be challenged by recent empirical work. Contrary to the prediction that anger motivates destructive, or retributive, collective action, a range of studies in the field of collective action have failed to establish a significant relation between anger and the motivation of such actions. Indeed, in complete opposition to this prediction, anger was observed to significantly motivate constructive actions. The key notion of moderators has been introduced and I have highlighted a moderator that is likely to play a key role in determining when anger motivates constructive actions.

---

[59] Indeed, the canonical papers that espouse a dual-pathway view of collective action (see van Zomeren et al. 2004), took both the anger and the instrumental reasoning pathway to motivate constructive actions in equal measure. Tausch (2011) then sought to test this prediction explicitly and gathered the confirming results discussed above.
[60] I will return to this in section 3.4.

Proponents of the *inefficacy claim* take the communication of anger to be ineffective in ameliorating the predicament of oppressed groups largely because their communication of anger will only serve to antagonize the dominant group. An antagonized group is one that is likely to retaliate against one's in-group, or at least avoid this group, and will therefore not be willing to work towards rectifying injustice. Research in intergroup psychology provides mounting evidence against this, however, as communications of anger have been shown to correlate with increased support for constructive actions tendencies on behalf of the dominant group, towards those expressing anger.

One experiment, for example, probed the effect of anger communication on the responses of Americans to Syrian-American relations. In the experiment, Americans watched a short video clip about Syrian-American relations after reading a brief text. In the 'anger' condition, the text described how a key Syrian leader gave an enraged speech that was aggressive towards US. In a 'hope' condition, the text described the leader's hopeful view on the resolution of the conflict. And finally, in the neutral condition, non-emotional factual information was relayed in the text about the Syrian leader's speech. American subjects were then asked to register their support for conciliatory policies, such as continuing exports of food and medicine to Syria, and accepting Syria's request for the US to fund humanitarian projects in Syria (Tagar et al. 2011).

The Out-group Reason for holding the *inefficacy claim* would predict Americans to become antagonized by, and respond retributively to, displays and communications of anger. Nussbaum's view would, therefore, predict support for conciliatory policies to be lowest in the anger condition. The view would arguably expect increased support for conciliatory policies in the neutral control condition because participants would be able to think clearly about the conflict at hand, and not be negatively biased by anger. Similarly, the view would likely expect increased support for conciliatory policies in the hope condition as well.

Contrary to these predictions, however, the study found support for conciliatory policies to be highest in the anger condition. Support for conciliatory policies was not only higher in the anger condition than in the control condition, but the anger condition even saw significantly higher levels of support for conciliatory policies than were observed in the hope

condition. This starkly opposes the predictions we would expect of the *inefficacy claim*, and suggests that anger communication has an important role to play in inter-group conflict resolution. Mounting evidence supports the main finding of this study, as anger communication has been observed to increase dominant group support for conciliatory policies in the context of the Israel-Palestine conflict, US race-relations (Shuman et al. 2018), and cases of xenophobia (de Vos et al. 2013). The Out-group Reason for holding the *inefficacy claim* is therefore challenged by evidence that the communication of anger from a disadvantaged group often actually increases dominant group support for conciliatory policies.

In one of the above-mentioned studies, researchers investigated how the communication of anger plays such beneficial roles in inter-group conflict. One might think that perhaps it is fear of the enraged group that causes the increase in support for conciliatory policies on behalf of the dominant group. Indeed, Nussbaum notes that in so far as anger can act as an effective deterrent to keep others from infringing upon one's rights, it does so by inspiring fear which is 'not likely to lead to a future of stability or peace' (2015: 55). Contrary to this, however, increased levels of empathy were observed in the dominant group following anger communication from an oppressed group (de Vos et al. 2013). Indeed, empathy for the oppressed group was actually significantly higher in the anger communication condition than in the control condition where the unfairness of the situation was communicated unemotionally. This suggests that anger communication is a potentially optimal way of recruiting the empathy of dominant group members that are the target of the anger, and that this in turn mediates their increase in support for conciliatory policies[61]. This suggests that empathy for, rather than fear of, the oppressed group may cause the positive effects of anger observed in other studies.

In a follow-up study, de Vos et al. (2016) found the appropriateness of the anger to moderate its positive effects. In other words, anger communication increased empathy in the dominant group when the dominant group saw the oppressed group's anger as a justified response to the situation at hand[62]. This highlights a key moderator that helps determine when anger communication is likely to cause the out-group to support constructive policies.

---

[61] Mediators, in psychology, are variables that speak to how or why certain effects occur. In this case, empathy is a mediator because it can be seen as an intermediary step that explains the effect of anger on outgroup support for constructive actions.
[62] I discuss this further in section 3.4.

In sum, we first saw that empirical work supports a vital role for anger in motivating constructive collective action. In addition to this, studies also showed the communication of anger on behalf of those oppressed to trigger dominant group support for constructive and conciliatory actions. Both reasons for holding the *inefficacy claim* are therefore challenged by recent empirical work.

Against any charge of having cherry picked the experimental work I rely upon, it is crucial to note that the trend in the experimental literatures I invoke have departed from debates on whether anger is beneficial, or not, to collective action or inter-group disputes. Instead, they focus on trying to uncover key moderators that determine *when* anger is destructive or constructive, so as to better understand and aid conflicts. The underlying commitment to anger's pluripotency is clear and widespread (Spring et al. 2018). This should be reflected in contemporary philosophical treatments of anger and the relevance of particular moderators should be attended to.

The above considerations suggest that we should drop the *inefficacy claim*, at least regarding anger's role in inter-group conflict resolution. If we do so, however, what does this mean for the *nature claim*? I turn to this question now.

3.3   Anger's Nature Revisited

The question that arises once we take Nussbaum's *inefficacy claim* to be misguided, or, at the very least, overly simplistic, is whether we should do away with the traditional take on the nature of anger altogether. We seem to have two types of choices here. Either we stick with the traditional *nature claim*, and take anger's oft constructive role in fighting for justice to be an exception to anger's typically punitive nature, in which case an explanation must be given for why anger behaves uncharacteristically in the social justice case; or, we take anger's constructive role in social justice to be evidence of something important about the nature of anger itself. What exactly this might be must be cashed out, but any account of this sort will involve a rejection, or at the very least a modification, of the *nature claim*. I will now argue that we should pursue an option of the second, rather than the first, type. After this, I sketch a positive proposal of the second type.

One option open to anyone wishing to secure a fundamental role for anger in the fight against social injustice, is to do so by distancing anger in these cases from cases of everyday anger. Doing so involves casting anger's constructive role in fighting social injustice as a special, or deviant, case that departs from anger's nature. Interpersonal anger, i.e. anger felt for one person by another, is taken to be the paradigmatic case of everyday anger. In line with the *nature claim*, everyday interpersonal anger is, on Nussbaum's view, ineffective at bringing about interpersonal resolutions (2016: Ch.4). By constitutively involving an aim for payback, everyday interpersonal anger may prompt retributive behavior that escalates the dispute, proving counterproductive for the angry party. Everyday anger typically has an individual to blame and enact payback over. Anger in the case of social injustice, however, might be considered less destructive in those cases where payback cannot be exacted against a particular person or group as easily. This is, perhaps, because there is often no adequate individual to blame in these scenarios, and hence payback is a less immediate concern than real change.

Rosen (in progress) and Swaine (1996) have made independent cases for the constructive role of anger in social justice, that hinges on anger in these cases being atypical for its lack of a clear agent(s) to blame. That is, as the object of anger is not typically an individual in such cases, payback cannot be straightforwardly exacted. Such a view maintains a commitment to the *nature claim,* as anger is still by nature punitive. Cases of constructive everyday anger are seen as outliers, and the constructive role granted of anger in promoting social justice, is taken to hinge on a sort of fluke in the natural functioning of anger; that anger is not able to live out its natural function of procuring payback or retribution.

I would like to argue for an alternative to this story; one that secures anger a constructive role in the fight against social injustice, in line with, rather than despite, its nature. This will be a sketch, but one that I think proves more attractive than the other options on the table. I would like to start by briefly arguing against the dire picture painted of everyday anger, present in any attempt to capture the constructive role of anger in social justice as an 'atypical case' of anger. The next section challenges the view that anger is typically ineffective at resolving interpersonal conflicts.

3.3.1 Interpersonal Anger

I have been discussing anger in cases of social injustice. One way of accounting for anger's constructive effects in this domain, is to take anger's efficacy in this domain to hinge on it not having clear objects to enact retribution over (Swaine 1996; Rosen, in progress). If this sort of view gets things right, then the possibility of split referents re-emerges as a potential resolution to the debate on anger. That is, such views would support one type of anger being constructive, namely anger in cases of social injustice, while condemning other types of anger as typically destructive. The dispute over whether to condemn or praise anger would then dissolve into which anger is being referred to. In this section I will argue that this type of view endorses an implausible account of anger's nature. Discarding these views will, in turn, close the door on split-referent attempts at dissolving the anger debate.

To evaluative the claim that anger's retributive tendencies depend on the type of object anger is felt toward, we should get clearer on what exactly the objects of anger is thought to be directed at across the relevant cases. We saw, in chapter 2, that emotions can have two types of particular object: targets and propositions. Particular objects are distinguished from formal objects, which are the evaluative properties attributed to, or instantiated in, the particular objects of the emotion. The formal object is taken to be constant across emotion types, so we will take it to be constant across anger in cases of social injustice and interpersonal anger. How does the particular object of anger differ between these two cases however? The particular object of anger in cases of social injustice vary. Anger in cases of social injustice can have as its target; groups (such as an ethnic group), social objects (such as institutions) and individuals (that have committed or perpetuated a social injustice). In addition to these, anger in social injustice cases can lack targets but have propositions, regarding the state of (unjust) affairs, as their particular objects (that there be structural gender oppression, for example). Interpersonal anger, on the other hand, always has an individual as its target, and can admit of a variety of propositional objects. Interpersonal anger and anger in the social injustice case, therefore, overlap, when the target of anger is an individual, and the propositional object involves an unjust state of affairs to which the individual is causally related.

The claim that anger in cases of social injustice do not have clear objects over which to enact retribution is clearest in cases where anger lacks a target, and has only a state of affairs, or

proposition, as it's particular object. In these cases, there doesn't seem to be a target for retribution. In cases where anger has as its target social objects, or groups, there is an entity that is blamed, but, the thought is, these targets are less straightforward to enact retribution over than individuals. With less straightforward targets for retribution, anger in cases of social injustice will be more effective and constructive than interpersonal anger, according to this view (Swaine 1996; Rosen, in progress). We will see that we have reason not to endorse such a view.

By 'interpersonal anger', then, I will mean cases of anger that have individuals as their target. These are to be distinguished from cases of anger at social injustice, by which I will mean cases of anger that have groups or social objects as their target objects, and/or states of affair as their propositional objects. I will exclude cases of interpersonal anger that overlap with social justice cases for simplicity, as well as for a few more substantive reasons. Firstly, doing so seems to capture the relevant contrast in cases between those where anger is meant to have clear targets for retribution, and those in which it is not. Secondly, in doing so, I remain in line with the use of the term 'interpersonal anger' in the empirical literature. The experimental paradigms used in mainstream and canonical work on 'interpersonal' anger involve personal slights and situational conflicts, as well as unfair economic distributions, rather than social injustices, as their stimuli. I do not deny that there are cases of anger that target individuals for social injustices, indeed I take this to be a very prevalent form of anger, but for our present purposes in assessing the differences between interpersonal anger and anger in cases of social injustice, it is best to proceed with paradigmatic conceptions of each, which I take myself to do be doing. We will see that despite having clear targets of blame, over which retribution is straightforward to enact, interpersonal anger is not as retributive as has been assumed.

Many take anger to be ineffective at resolving interpersonal conflicts for reasons analogous to the In-group and Out-group Reasons discussed above of the social justice case (see Pettigrove 2012; Nussbaum 2016: Ch.4). That is, that anger motivates destructive actions against the person one is angry at, and that acting on, or communicating, anger is likely to be met with animosity or retaliation from its recipient. I argue against this simplistic view of anger's role in interpersonal conflicts, and in so doing highlight similar moderators to those discussed in section 2, to be relevant to the role anger plays in interpersonal conflicts.

Experiencing interpersonal anger is very common. Indeed, it has been observed that most people become 'mildly to moderately angry anywhere from several times a day to several times a week' (Averill 1983: 1146). It is safe to say that we experience anger far more frequently than we are involved in interpersonal conflict. Often, when one communicates one's anger, one is met with compassion, apology, and understanding. Other times, one is not, and this may intensify one's anger and lead it to be expressed in retributive ways, but our everyday anger does not necessarily, or perhaps even typically, involve retributive actions.

When angry at loved ones, communications of anger do not invariably lead to escalations in conflict, they often lead to new levels of understanding and intimacy. Even when angry at a stranger, anger can be non-retributive. When someone cuts in front of you in a queue at a coffee shop, for example, you may justifiably become angry. In such a scenario, you want them to realize that what they have done is wrong, that you won't accept it, and that they are mistaken in taking themselves to be entitled to such behaviour. At least in some cases, tapping them on the back and getting them to join the end of the queue would satisfy your anger. My point is not just that aggression is far less typical than are everyday occurrences of anger, but that the goal of your anger, often doesn't seem to be in any sense retributive. You don't want the queue-cutter to be denied coffee, nor do you want something bad to happen to them, rather, you just want them not to get away with queue-cutting. To enforce this norm is not to punish, or get pay-back, over the queue-cutter, but merely to get them to stop committing a wrong.

Much like anger in the social justice case then, interpersonal anger seems to similarly pick up on wrongs, and the communication of anger between individuals seems to at least sometimes be a crucial means for change and understanding. Indeed, empirical work challenges the In-group and Out-group Reasons for holding interpersonal anger ineffective. That is, empirical evidence challenges the predictions that interpersonal anger typically motivates destructive behaviour, and that it tends to antagonize its recipient. With regards to the types of action interpersonal anger motivates, a canonical study by Averill (1983) found a higher percentage of non-aggressive than aggressive action tendencies in people experiencing anger. Although there is robust evidence that, on economic distribution paradigms[63], angered individuals *on*

---

[63] Economic distribution paradigms typically involve engaging participants in a (often virtual) two-player game, or task, where the payoffs for each player depend on the simultaneous choice of both players. The task creates incentives to cooperate as well as to exploit or punish one's partner player. Prisoner Dilemma tasks are paradigmatic examples of this type of paradigm. These paradigms involve many repeated trials so as to study the development and violation of trust and cooperative behaviour.

average respond more punitively to unfair economic distributions, than to fair economic distributions, recent work has observed that angered individuals still choose behaviours that are economically *cooperative* in response to unfair economic distributions (Klimecki, et al., 2018).

This suggests that anger may not typically motivate retributive actions towards the individuals it is directed at. Which actions anger motivates is likely to be more a question of context than object. Indeed, much like the social justice case, the ability of the target of anger to change may be key to determining whether anger motivates constructive behaviour or payback oriented behaviour. A study on adolescent responses to bullying for example, found implicit beliefs about bullies to predict desires for revenge (Yeager et al. 2011). Greater desire for vengeance was observed in participants who held implicit beliefs about the fixed character traits of the bully. The reverse was found for participants that held implicit beliefs about the changeability of the bully's character. By highlighting the changeability of the target of vengeance as a moderator, the interpersonal case seems far more similar to the empirical work I surveyed regarding the social justice case.

Similarly, as opposed to antagonizing the object of anger and promoting retaliatory behavior on their behalf, the communication of anger in interpersonal contexts has been found to have constructive effects. The communication of anger has been observed to trigger increased social support in close relationships (Yoo et al. 2011), as well as increase personal gains in interpersonal financial negotiations (Van Kleef and Côté 2007). The latter study pinpointed the appropriateness of anger as a key determinant of anger's beneficial effects in interpersonal negotiations. The constructive effects of anger were highest when the anger was seen as justified, as the offender compensated the low status party in these cases. Indeed, they found that when the low status individual was unjustified in his anger, the offender penalized the angry individual most harshly (Van Kleef and Côté 2007). This points to the crucial role of the appropriateness of anger in moderating whether the anger is well received or not. This again suggests that interpersonal anger is far closer to anger in the social justice case; as appropriateness similarly acts as a moderator over anger's constructive effects.

We are left with little reason to think anger in interpersonal contexts is more destructive than anger in the case of social injustice. The studies above suggest interpersonal anger can play robust beneficial roles; by motivating non-aggressive actions, being well received by its target

and bettering the angry person's situation. The constructive effects of interpersonal anger also seem to be moderated by analogous factors seen to be key in the social justice case (sec. 3.2), such as appropriateness and the changeability of the object of anger.

If everyday anger enjoys constructive consequences under certain conditions, as the empirical work suggests, and our everyday experience of anger is not typically retributive, as I have suggested, then a view that singles out anger in the social injustice case as atypically constructive, seems unfounded. Appealing to the type of object anger is felt towards, seems, therefore, not to be a viable feature on which to hinge one's praise or condemnation of anger. Given that interpersonal anger, the paradigmatic case of having a target object to blame and enact retribution over, can be effective in interpersonal conflict, the constructive role of anger in cases of social injustice doesn't seem to hinge on lacking straightforward targets. In other words, the object of anger doesn't seem to be the feature on which anger's constructiveness depends. In the next section I sketch an alternative view of anger and its efficacy that better fits the experimental evidence discussed so far.

3.4 Sketching an Alternative

We have seen that we have much reason to reject a traditional view, such as Nussbaum's (2015; 2016), that takes anger to be constitutively tied to payback. We therefore have reason to rethink the *nature claim*. I take the empirical work discussed in this chapter and the last to suggest that the *nature claim* is unlikely to be true of anger. Specifically, I take there to be two at least distinct pathways through which anger can pursue change, and I take there to be certain moderators or factors that are key to determining which pathway is pursued. I call these two distinct anger pathways the Retributive and the Recognition Pathway, respectively. I take both these pathways to be characteristic of anger, whether it is interpersonal anger or anger at social injustices. I will therefore have a few things to say about anger in general before moving on to what my proposal has to say about anger in the social justice case. I will begin with two examples that will help characterize the distinction between the proposed pathways.

3.4.1 Recognition and Retribution

Imagine your friend Miguel manages somehow to steal your inheritance. You would surely want to make him suffer for committing this wrong against you. You may make him suffer by cutting him out of your social circle, suing him, defaming his character or threatening him with physical violence. Your anger would, therefore, follow the Retributive Pathway and be aimed at Miguel's punishment. Let's suppose now that you fail to act on your anger. This could occur for a range of reasons, perhaps there was high social pressure not to act, or perhaps you merely lacked the opportunity. The desire to make Miguel suffer is still key to your anger in such cases. Indeed, you are likely to hope for things to go badly for Miguel, and become happy upon hearing about his own hardship, even if this news comes years later and is entirely unrelated to you. Your desire for payback is satisfied in hearing this news, and you see his suffering as deserved.

Now imagine you are angry at your mother for not being there for you throughout your divorce. When you needed her support the most, your mother decided to go on a spontaneous three-month long holiday abroad. Your anger at your mother would be entirely justified, but it is unlikely that you would wish to ensure her social exclusion, or defame her character. Nor would you want to make her suffer physically, or hope for things to go badly for her in the future. Your anger's goal does not seem to be that your mother suffer, but rather, that she understand what she has done. Whereas Miguel's suffering satisfied your anger's desire for payback, your mother's suffering will not satisfy your anger in this case. This is because your anger at your mother does not involve a desire for payback but rather for recognition. Your anger's desire for recognition will be satisfied by your mother's genuine acknowledgement of the wrong she has committed against you. This will involve your mother sharing your appraisal of her actions towards you as unjust. This case of anger proceeds along the Recognition Pathway as it aims for an epistemic change in the offender.

These examples illustrate the key differences between my two proposed pathways. I take the Retributive and the Recognition Pathways to both be typical of anger. Each involves a distinct desire, payback and recognition respectively, and associated satisfaction conditions, suffering and understanding respectively. Each pathway also involves its own set of typical actions, ones that aim to satisfy the respective desire. Explicit communications of anger, such as conversation or formal writing, will be characteristic of the Recognition Pathway while

retributive behaviour, from physical violence or property damage to subtle forms of payback, will be common of the Retributive Pathway. I take Nussbaum to equate anger to what I have characterized as its Retributive Pathway, while at best underestimating, and at worst, outright denying, the Recognition Pathway of anger.

In appraising a situation as unjust, anger appraises things as in need of change. Anger can be said to involve a general desire that particular wrongs or offenses not continue or repeat themselves. I propose that this general desire for change becomes a desire for retribution or recognition respectively, depending on the circumstance and context. The unjust circumstances can be seen as propositional objects of anger, or reasons for being angry, while the agents that are blamed for them are the targets of anger. The Recognition Pathway aims to change the reasons for anger by getting the target of anger to share in their appraisal of the situation as unjust. The Retributive Pathway aims to change these reasons by exerting force over the target of anger and making them suffer. The latter says something like, 'respect us or we will smash you', while the former says, 'you are disrespecting us and this is wrong'. The Retributive Pathway attempts to bend others to one's will, while the Recognition Pathway attempts to have its moral appraisals shared by the targets of anger.

A worry may arise regarding the plausibility of anger-motivated actions falling neatly within one pathway, or the other. Shouting at someone who wronged you, for example, could be seen as falling into both pathways, as this can be both communicative and punitive. This ambiguity is not worrisome, however. The two pathways I propose are not equated to the action tendencies that typically accompany them. It is the desire associated to the anger that characterizes which anger pathway is being pursued, rather than the actions themselves. There might seem to be another worry over which desire can be said to be involved in each instance of anger. That is, often anger can cause both epistemic as well as punitive consequences on its targets, such that it might seem that both a desire for payback and one for recognition would be satisfied. I agree that this is often the case, but I contend that in each instance of anger only one of these desires is typically primary. For example, in the case of your friend Miguel, perhaps making him suffer will lead him to come to understand how wrong he was to have stolen your inheritance. But this is not your anger's goal, it is just an unintended consequence of your actions. Indeed, even if Miguel acknowledged that what he has done is wrong, this would not satisfy your desire as you could still wish to make him suffer for his actions. So, epistemic change in this case is tangential to your desire to make

him suffer. In the case of anger at your mother, the reverse is true. That is, whereas your anger's goal is that she appraise her own actions as wrong, this realization may be an emotionally painful experience for her to undergo. But this suffering is an often undesired consequence of your actions achieving their desired goal, which may be entirely epistemic.

Indeed, if we briefly consider more closely the workings of apology, the robustness of the Recognition Pathway should come to the forefront. Apology can be seen as one of the main satisfaction conditions of anger's desire for recognition. By this I mean that genuine apology is a sufficient satisfaction condition for the desire for recognition. The apology must, of course, be sincere. It must reflect the genuine epistemic state of the offending party. This does not mean that apology erases the reasons for anger. Apology reflects the offending party's evaluative stance towards these reasons. What apology seems to do then, is signal to the aggrieved party that the offence should not have occurred, that they know it was wrong, and that they will at least try to avoid committing such an offence again. Apology is, therefore, in a sense both past-oriented and future-oriented. As is anger. Anger is a stance on offences that have occurred (or are occurring) and involves a desire for these offences not to continue or repeat themselves. I have suggested that this general desire for things to change, or not repeat themselves, in anger, can manifest in two distinct desires: for retribution or for recognition. That anger often terminates when a genuine apology is issued should be taken as evidence for the robustness of the Recognition Pathway. Genuine apologies give one reason to revise one's anger, not because the wrongs that triggered it have been excused or erased, but because one's desire for the wrong to be recognized as a wrong has been satisfied[64].

Sometimes anger meets its end even when an apology from the offender isn't given. In situations that might be called cases of vicarious apology, members of one's community recognize the wrong that has been committed against you. These agents need not issue explicit apologies, although they often do. When anger is met with the understanding of your community, it often comes to an end precisely because the desire for recognition of the wrongs committed against oneself have been acknowledged. One's status as an individual that should not be treated in the ways that triggered anger is recognized by others, and the wrong one has suffered is acknowledged as a genuine wrong by others. That anger sometimes meets its end in the acknowledgement of others gives us reason to think it often involves an

---

[64] For an account of forgiveness that involves a conception of anger in line with my own see Hieronymi (2001)

aim for recognition, for we often express our anger to members of our community from whom we do not seek any type of retribution or payback. If anger's aim were always retributive it would be immune to vicarious apology, for in cases of vicarious apology there is no suffering or status-lowering in those recognizing the anger as justified. When the apology is issued by the offender, it typically comes with regret, guilt, or some milder form of psychological pain. This makes it harder to see that anger's aim was for recognition rather than punition. Cases of what I have called vicarious apology avoid this worry and allow recognition to emerge as a genuine satisfaction condition for anger.

3.4.2 Anger at Social Injustice

I will now focus on how this framework accounts for anger's constructive role in fighting social injustice. Outlining the Recognition Pathway of anger helps make sense of the experimental evidence regarding anger's role in collective action and intergroup conflict. These studies showed the *inefficacy claim* not to hold for many cases of anger. These are ones, which the Recognition Pathway nicely captures, where anger was observed to motivate communicative actions such as protesting, petitioning, and lobbying, rather than punitive actions. The In-group Reason for the *inefficacy claim* does not hold for the Recognition Pathway because by aiming at recognition rather than harm, communication rather than payback is central. In terms of the Out-group Reason, I contend that the dominant group can understand anger as either an appraisal of injustice looking to be shared, or as a wish for payback. The former will favour the Recognition Pathway as the dominant group will tend to empathize with the oppressed group and support conciliatory policies towards them. The latter will favour the Retributive Pathway, as the dominant group will be more likely to pursue retaliatory actions against the oppressed group, or withdraw from any engagement with them. The Out-group Reason for the *inefficacy claim* can, therefore, be seen to hold for the Retributive but not the Recognition pathway.

I take the moderators highlighted in section 3.2 to be important for at least two reasons. Firstly, they help make sense of when each anger pathway will be pursued, by providing key contextual features which moderate over a) the desires for payback or recognition of the angry, and b) the reaction (support or retaliation) of the group targeted by anger. The moderators therefore play a crucial role in determining which pathway will be at play in a

given occurrence of anger. Secondly, they help explain why the traditional view of anger has been so prevalent. I will discuss these two points in succession.

The two moderators crucial to determining which of the proposed anger pathways is favoured are: the changeability of the target of anger and the appropriateness of anger. These were shown to moderate the effects of angry behaviour and anger communication respectively (sec. 3.2). When the target of anger is changeable, the Recognition Pathway tends to be pursued, as the target may seem more open to sharing the appraisal of the situation as unjust. When the target of anger is seen as unchangeable, however, the Retributive Pathway is more likely to be favoured. Indeed, recall that the experimental work indicated that aggressive actions were most typically pursued when the object of anger seemed entirely unalterable, in 'nothing to lose' scenarios.

The appropriateness of anger, on the other hand, was key to determining how anger communication was received by the target of anger. When anger was seen as appropriate, the Recognition Pathway can be seen to be favoured, as the dominant group supported the oppressed counterpart and engaged in constructive behaviour towards them. The reverse was true for anger seen as inappropriate, which favored the Retributive Pathway by making retaliation or disengagement with the angry party more likely.

The empirical work, therefore, suggests that the Retributive Pathway is most likely to occur when the target of anger is seen as unchangeable and the anger of the oppressed as inappropriate. This is actually an apt description of the state of affairs in severely unjust societies. Power imbalances can structure which instances of anger are seen as appropriate or inappropriate, such that the status quo is perpetuated and injustices left unaddressed (see for example Frye 1983: 84-94). We saw this to be supported by historical work in chapter 1, where the anger of peasants is deemed irrational, while that of King's righteous and appropriate. This means that the anger of the oppressed will more easily be dismissed as inappropriate due to dominant ideology and bias. Relatedly, the angry are less likely to take their targets to be changeable in real life cases of entrenched social injustice, given their lived history of struggle, and are therefore more likely to act destructively. The empirical work surveyed in this chapter therefore stands in line with plausible sociological observations of life under oppression. But, of course, moderators aren't binary, and, therefore, these moderators will be present to greater or lesser extents in oppressive societies. This suggests

that the Retributive Pathway is made more likely under conditions of severe social injustice, where both moderators are deeply engrained features of that society.

My account has the resources to make sense of when anger is likely to involve desires for retribution, without taking anger to be constitutively tied to a desire for retribution. It, therefore, seems superior to the traditional account of anger as constitutively involving a desire for retribution. Such a traditional view sits uncomfortably with the empirical work I have highlighted, and entirely neglects what I call the Recognition Pathway. Does Nussbaum, and others who endorse the *inefficacy claim*, get things right regarding anger under conditions of oppression, however? The *inefficacy claim* seems to get things partly right, in so far as it provides a description of how I have characterized anger to operate under conditions of severe social injustice, namely in destructive and retributive ways that invite retaliation. But, as a general account of anger's efficacy under conditions of oppression, it is incomplete. We have reason to think that, despite moderators that favour the Retributive Pathway being more prevalent under conditions of social injustice, anger is still an extremely effective way of constructively fighting against social injustice.

For example, although under conditions of social injustice the dominant group might tend to dismiss the anger of the oppressed as inappropriate, there are many cases of dominant group members becoming allies of oppressed groups under conditions of severe oppression[65]. From the studies surveyed in section 3.2 above, we have reason to think that anger is one of the most effective ways of recruiting allies, as anger was seen to recruit more support from out-group members than neutral communications of wrongdoing (Tagar et al. 2011). Relatedly, studies have shown that having even just *one* individual member of an out-group share the in-group's anger, results in the in-group seeing the out-group as potential allies. This in turn correlates with in-group support for conciliatory actions towards the outgroup (see McDonald et al. 2017). This highlights the crucial effect of recruiting allies under conditions of social injustice. Recruiting even one single member of the outgroup seems to impact the perceived changeability of the out-group immensely, and increased changeability was one of the moderators highlighted that favoured the Recognition Pathway.

This suggests that even amidst potential widespread dismissal of the anger of the oppressed, securing even a few out-group allies is not only possible, but perhaps most effectively

---

[65] See Brown (2002) on white allies to the US civil rights movement.

achieved through displays of anger. Individual allies can be sensitive to the epistemic value of in-group anger, and this can lead to changes in the key moderators such that the Recognition Pathway is further favoured. This suggests that anger has a robustly effective and constructive role to play in struggles against social injustice. What seems plausible, is that the Retributive Pathway may be favoured mostly in 'nothing to lose' scenarios, where attempts at recruiting allies have proved futile, and one's anger has been systematically dismissed as inappropriate. 'Nothing to lose' situations will be ones where the two moderators indicated in the empirical literature – perceptions of appropriateness and perceptions of liability to change – are most clearly operative.

This sheds some light on the reasons for why the traditional view of anger is so prevalent. It is plausible that anger will have commonly manifested itself as retributive in unjust societal arrangements, where the two moderators highlighted would have been operative. This may have led to retribution being viewed as part of its nature. Moreover, it would have been in the interest of those in power to dismiss anger as inappropriate and to perpetuate the view that anger is destructive. One of our key moderators was perceptions of appropriateness. Therefore, if those in power deemed anger inappropriate, this moderator would be intensified, which in turn would favour retributive actions from those in anger. This, in effect, would 'prove' the traditional construal of anger correct, much to the benefit of those with a vested interest in maintaining the prevailing status quo. Once anger's retributory tendencies are understood as dependent on specific features of the very injustices it seeks to combat, rather than understood as constitutive of anger's normal function, we have prima face reason to take the anger of the oppressed seriously and to acknowledge its efficacy. A view of anger committed to the *nature* and *inefficacy claims*, then, is guilty of reading into the very nature of anger, what on my account are contingent tendencies of anger in 'nothing to lose' scenarios. In doing so, the traditional view leaves little room for the constructive potential of anger to be theorized, or acknowledged in practice.

I have given the traditional construal of anger a limited, but important, role in my account. The traditional view of anger as constitutively tied to retribution obscures the robust constructive role anger can play in struggles for social justice. Anger, on my account, involves a desire for retribution under particular conditions, particularly when the angry agent has exhausted other opportunities, as these will be cases where the key moderators are deeply entrenched. In so far as Nussbaum captures the Retributive Pathway effectively then, she

does so by obscuring the social dynamics on which, according to my account, it depends. By taking anger's retributory effects to go hand in hand with its nature, the dependence of the *inefficacy claim* on the social contingencies that make it sometimes the case, is concealed. My view gives us the resources necessary to see Nussbaum's partial descriptive success to hinge on features of extreme unjust social situations, rather than on anger's nature itself. In doing so it allows us to see anger as a prima facie legitimate state to express and act upon in pursuing social change, as well as one that can be effective in constructively enacting this change. My view accounts for the relevant experimental work within a framework that secures a more plausible view of the nature of anger. In developing it, I have built a head on challenge to inefficacy claims and leant support to the view that recognition is among the main aims of anger.

3.5 The First Two Desiderata

In this chapter and the last we have made significant steps towards meeting the first two of our desiderata; the provision of an account of anger's nature, and of an account of anger's instrumental rationality. We have challenged the dominant conception of anger's nature as constitutively tied to retribution and we have seen that anger can be effective in constructively fighting against social injustice. I discuss the meeting of each of these desiderata in turn below.

3.5.1 What is Anger?

The first desideratum of our foundational approach was to provide an account of the nature of anger. Doing so was expected to move the debate forward on a number of fronts.
Firstly, that of referent. There have been two points at which the anger debate could have been dissolved by establishing that each side of the debate is actually referring to a distinct phenomenon. The distinction between basic and non-basic anger could have supported this view, whereby those who condemn anger have basic anger in mind, while those who praise anger have non-basic anger in mind. We saw in chapter 2 that this distinction is not supported by our evolutionary history, nor by contemporary neuroscience. The picture of human anger referring to a modular affect program on the one hand, and a cognitive non-basic form of anger on the other, was seriously challenged. We saw that anger is best

construed as heterogeneously constructed by a mixture of biological and cultural inheritance that, crucially, reflects adaptive pressures for cooperative living. The debate on anger cannot be dissolved by appeal to distinct referents on this account.

The second point at which a mismatch in referents could have been argued for, involved taking anger's objects to individuate robustly distinct types of anger. The thought would be that those who condemn anger are referring to anger aimed at objects over which pay-back can be exacted, while feminist philosophers, on the other hand, are referring to anger directed as objects over which the exaction of pay-back is far from straightforward. One initial reason to reject this move is that some of the thinkers that have most forcefully condemned anger, have taken themselves to be referring to the same kind when discussing anger in interpersonal and social justice cases (Pettigrove 2012; Nussbaum 2016). I have argued that the move to hinge one's condemnation of anger on how straightforward it is to exert pay-back over their objects is unsuccessful. This is because such a move casts interpersonal anger as paradigmatically retributive, a contention that I have challenged.

In section 3.3.1 I highlighted the similarities between anger in the interpersonal and social justice cases. In both cases, anger was seen to often play constructive non-retributive roles, and the same contextual features that moderate over when anger pursues the Recognition Pathway, or the Retributive Pathway, in the social justice case, were seen to be at play in interpersonal cases of anger as well. This gives us reason to doubt a picture where interpersonal anger is seen as more retributive than anger in the social justice cases.

We have, therefore, seen reason to reject attempts to dissolve the anger debate by appeal to distinct referents. We can, however, argue that those who condemn anger are committed to a simplistic and mistaken view of the nature of the emotion, and indeed this is what I have been doing. Because the anger debate has not benefited from a systematic foundational approach, detailed accounts of anger's nature have not been provided by either side of the debate. We can infer some commitments regarding the nature of anger from each side of the debate. Those who condemn anger are likely committed to a modular view of anger as a hardwired affect program, which I argued in chapter 2 should be rejected. Those who condemn anger have been explicit about their commitment to the claim that anger constitutively involves a desire for retribution. Chapter 2 also spoke against this, by highlighting the difficulty in pin-pointing a single evolutionary function for anger, especially

once our adaptation into collaborative living is taken seriously. In this chapter I have challenged the retributive reading of anger's nature head on, focusing on the particular context of social injustice. I have argued that anger is best seen as involving two distinct desires, where contextual moderators predispose their occurrence.

Anger is perhaps best seen as involving the open-ended desire to change the situation that triggered it[66]. The particular pathway pursued in order to satisfy this desire for change will, at times involve retribution and, at others, recognition. Attempts to assimilate recognition to a form of retribution should be resisted for a number of reasons. Firstly, given the empirical evidence canvassed in this chapter and the last, we have reason not to prioritize a retributive reading of anger. Anger was been seen to play a number distinct roles, both in the context of social injustice, as well as in cooperative non-human species, such that a retributive function should not be held as primary. Furthermore, we have seen that the Retribution Pathway is favoured in 'nothing to lose' scenarios. The retributive view of anger is, therefore, at least sometimes, a self-fulfilling prophecy contingent upon features of oppressive societies. This suggests that retribution is significantly dependent on situational contingencies, rather than constitutively linked to anger. We therefore need strong reasons to convince us that what I have called the desire for recognition, is best seen as a subspecies of the desire for retribution.

Another reason to resist the assimilation of the desire for recognition to a form of retribution, is that this risks extending the term 'retribution' beyond meaning. If aiming for epistemic changes in other is considered a subtle form of retribution, we risk making the term so permissive that it lacks much explanatory power. Finally, it seems to me that one could argue in the other direction, and attempt to assimilate retribution to a form of recognition. The story would go like this: Anger constitutively involves a desire for recognition. That is, anger constitutively aims for one thing, that the target of one's anger acknowledge the harm or wrong they have committed. In cases where the agent has nothing to lose, this desire for recognition manifests itself in retributive and punitive actions. That is, anger only involves a desire for retribution when its desire for recognition cannot easily be met by other means,

---

[66] This means that anger involves both a cognitive component that represents things in the world to be a certain way, as well as a conative component for how the world should be. That anger, amongst other emotions, seems to have both traditional directions of fit, mind-to-word and world-to-mind, has been considered part of their distinctive or sui generis intentionality (see Scarantino 2010).

and agents therefore engage in acts that 'force' targets to recognize the reasons for their anger.

I have not argued for such a story, nor will I, but that such a story holds initial plausibility, should count as reason against pursuing the inverse strategy, namely attempting to cash the desire for recognition as a type of retribution. Our view of anger moving forward will be as follows; Anger is a phenomenologically salient intentional state triggered by offences, broadly construed. Given anger's multiple triggers a broad construal of 'offence' is in order. By 'offence' then, I mean any norm violation, or states of affair that are unfair or wrong, whether or not it involves a personal offence against the angry individual. Anger has evaluative phenomenal content where its objects are felt to instantiate this evaluative property. Anger implicitly represents its target object (when it has one) as being responsible for the offence. Anger involves a general desire[67] to change the situation that triggered it, or to prevent it from reoccurring. This desire typically arises in the form of one of two distinct desires: a desire for retribution, or a desire for recognition[68].

More will need to be said on how the general desire for change relates to the specific desires for recognition, or retribution. Are the latter two instrumental desires that are had for their role in achieving the non-instrumental desire of change? This picture seems to get things right in so far as retribution and recognition are often ways of securing, or at least procuring, change. Such a picture seems to provide an apt description of anger in cases of social injustice. In such cases anger's aim is for change, and this can be pursued through at least two distinct instrumental desires: for recognition, or retribution.

However, that recognition and retribution can sometimes be non-instrumental desires, makes their general classification along this dichotomy more complicated. We saw that sometimes anger comes to an end through vicarious apology. Anger can also presumably sometimes be satisfied by an apology, even though the situation is unchangeable, or likely to

---

[67] This desire could be cashed out in different ways without threatening the force of my suggestion. I am inclined to favour a holistic view of desire that incorporates dispositions to act to change the situation, as well as either, dispositions to feel pleasure if the situation changes, or, that it appear good to the subject that the situation change. Either of these non-action-based components will deliver on the idea that the desire involved in one's anger is likely to be satisfied if the situation that triggered it changes. What type of changes satisfy anger will depend on the particular case, but typically it will be some change related to the negation of the state of affairs that triggered it.

[68] I leave aside unconscious forms of anger, as well as forms of anger that have no object. Following the dominant trend in the philosophy of emotion literature I focus on occurrent intentional forms of anger.

repeat itself. Similarly, there might be cases where our anger does not cease, despite its triggering situation having changed, because the target of one's anger has not acknowledged their role in committing a wrong. Anger's satisfaction conditions are likely to be complex, and context-dependent. Sometimes change will be sufficient to satisfy anger's conative component, sometimes recognition or retribution will be, and at other times a combination of these conditions will be necessary for the satisfaction and cessation of anger. This makes the question of anger's conative component(s) less than straightforward, and suggests that a generic account of them, applicable to all instances of anger, might not be a realistic goal for an account of anger. My alternative account of anger allows multiple options for anger's satisfaction conditions which is, albeit less straightforward, nonetheless a positive outcome. This seems preferable to a formulaic conception of anger as constitutively tied to the desire for retribution. Note that I have not claimed the desires for recognition and retribution to be exhaustive of anger, indeed the general desire for change is itself a distinct desire. I have, however, characterized a robust conative component of anger that has been neglected: the desire for recognition. I have also told a story for how the conative components of anger seem to operate under conditions of social injustice.

By endeavoring to provide an account of anger's nature I have taken steps towards meeting my second desideratum, as well as my first. This is because anger's nature has been tied to dispositions to act, and these dispositions to act have implications for the instrumental rationality of the emotion.

3.5.2 Instrumental Rationality

I have argued that anger can be effective at fighting against social injustice. There is an underlying assumption, bred by the traditional view of anger as typically retributive, that anger can be effective in this regard only in so far as it motivates aggressive and destructive actions. Such actions can lead to quick and profound change, and there is room to argue for their being justified. This has not been my focus however. I have, instead, argued that there is much evidence of a neglected route for anger's efficacy, one that is distinct from a desire for retribution. Doing so has strengthened the case that anger in cases of social injustice, as well as in general, can be instrumentally rational: anger can further even the short-term aims of the agent, rather than compromise them.

I have argued that anger's instrumental rationality is often tied to the emotion's desire for recognition. It is important to highlight that there are multiple ways of understanding the instrumental rationality claim. I take the claim that anger furthers the agent's aims to be ambiguous between the following two meanings: 1. That anger leads agents to select actions conducive to achieving the ends set by anger's conative component(s). 2. That anger leads agents to select actions conducive to achieving their all-things-considered aim.

It's clear that anger is much more likely to violate instrumental rationality in the latter sense than in the former. My examples will focus primarily on cases where anger plausibly involves a desire for recognition. If you are a dinner guest at your partner's boss' home, and you become angered by the host's political views, it might be instrumentally rational in the first sense to communicate this anger publicly, while it could well be irrational in the second sense. By communicating your anger in this case, you are acting towards the end of having a wrong recognized, while you might be jeopardizing your partner's career prospects by disrespecting their boss. Furthering your partners career prospects might have been your all-things-consider aim, and in this sense, acting on your anger will be instrumentally irrational.

If your anger is aimed at making your partner's boss acknowledge the harm they are committing in espousing harmful political views, for example, but you do not act on your anger, you are not instrumentally rational in the first sense, though you might be in the second sense. This is because you do not take any actions to further the end of making your partner's boss aware of their problematic views, making you potentially irrational in the first sense. You are rational in the second sense because your action, in this case omission, is preserving your all-things-considered end of making a good impression on your partner's boss.

The typical charge against anger's instrumental rationality involves anger being instrumentally rational in the first sense at the expense of the second. That is, anger is taken to often undermine the agent's all-things-considered, or more long-term, goals by motivating actions that prioritize one of the agent's short-term goals. This is because anger is seen as prompting hasty actions, and bypassing deliberation that could lead to the prioritization of ones all-things-considered aims. It has not been my aim to deny that this sometimes happens,

only to deny the plausibility of the charge against anger's instrumental rationality, as a correct general depiction of the emotion's effects.

Two shifts in our conception of anger lead us to question the accuracy of the picture of anger as instrumentally irrational in the second sense. Firstly, we saw that anger is not as closely tied to hardwired action tendencies as has previously been thought. This means that anger is not likely to be motivationally modular and hence actions done out of anger are not to be seen as reflex-like or impulses. Actions done out of anger are, therefore, not likely to be entirely independent from the agent's wider concerns. This means that anger is not immune to wider considerations and will not prototypically manifest itself as opposed to these considerations.

Secondly, in moving away from the traditional conception of anger as constitutively involving an aim for retribution, the view that anger is typically instrumentally irrational in the second sense is further weakened. Aiming for retribution as opposed to recognition is far more likely to undermine long-term goals, seeing as the former can burn bridges while the latter is an attempt to construct, or maintain, bridges. In lending support for anger's aim for recognition then, anger's threat of undermining all-things-considered agential aims is lessened.

Of course, in real life situations it is far more complex to determine whether one is instrumentally irrational in the second sense. First of all, agents might not have clear all-things-considered aims in given situations, and secondly, experiencing anger may cause one to endorse a distinct all-things-considered aim. Return to the dinner party hosted by your partner's boss. If the boss is espousing racist views, your aim to further your partner's career may be trumped by your desire not to condone racism. Furthermore, self-attributions of instrumental irrationality occur after the fact. One regrets having (or not having) confronted one's partner's boss at the dinner party only after the dinner party, when one's all-things-considered aim may have shifted. There are, therefore, cases where agents will take their anger to have been instrumentally irrational in light of considerations that might not have been available, or actual, at the time when the agent acted out of anger.

The traditional view, exemplified by Nussbaum, that anger is counterproductive or ineffective at securing the agent's aims takes anger to be typically instrumentally irrational in both senses. Besides often undermining all-things-considered aims, by triggering rash

behaviour without taking wider concerns into account, anger, on the traditional view, also undermines the emotion's specific aims, by causing disengagement or retaliation in its targets. When you call out your partner's boss for racist views, you can be instrumentally irrational in both senses if your angry call-out causes defensiveness and dismissal, rather than understanding, on the boss' (and other guests') behalf. If, in angrily calling out your partner's boss, you have failed to achieve the aim of your anger, namely to get the wrong that triggered it acknowledged, you may have also undermined an all-things-considered aim to impress your partner's boss. If this is the case, then you are instrumentally irrational in both senses.

This chapter has provided reason to doubt anger is instrumentally irrational in the first sense, as well as the second. We focused on the types of action anger motivates, as well as the types of response anger issues in its targets, in cases of social injustice. We did so such as to evaluate whether anger typically undermines its goal for improving the unjust status quo. This is precisely the type of case where anger's instrumental benefit has been most forcefully highlighted by feminist philosophers, and against which Nussbaum (2015, 2016) argues. Against Nussbaum, I argued that anger can be robustly instrumentally rational in fights against social injustice, due to its connection to the neglected desire of recognition. As anger, in these cases, involves a desire to change the status quo, seeking change through desires for recognition is a key way in which anger can be instrumentally rational in both senses. Desiring recognition, rather than retribution, is more likely to motivate constructive, as oppose to destructive, actions, and is more likely to be met with empathy and support for change. This makes anger that involves a desire for recognition instrumentally rational in the first sense, as anger often succeeds in getting reasons for it acknowledged, and in promoting change, when it aims for recognition. Anger is instrumentally rational in the second sense because pursuing desires for recognition, as opposed to pursuing desires for retribution, is much less likely to undermine distinct goals the agent might have.

It is important to note that anger's desire for retribution can be instrumentally rational as well. Often cases of anger at social injustices are characterized by short-term putative instrumental irrationality, and long-term instrumentally rationality. Recall the Chicago riots of 1968, the damage done out of anger was extensive and affected the African American community disproportionately. But these riots were instrumental in the fight for racial justice in the United States. Such cases are ones that can be construed as instrumentally irrational in the first narrow sense, and instrumentally rational in the second. If the aim of the agents'

anger was to improve the situation of their social group, they failed as things actually got worse in the immediate aftermath of their expression of anger. But the anger was instrumentally rational in achieving an all-things-considered aim of racial equality. Even when anger can be counterproductive, or instrumentally irrational, in the first sense then, it can often turn out to be instrumentally rational in the second. There is, however, a different way of reading such cases that is worth mentioning, where anger comes out as instrumentally rational in both senses. If anger aims for retribution because certain moderators are at play, that make the situation one in which the angry agent has little to lose, then anger can be rational in the first sense as well, for it leads to actions that are the only way of achieving the agent's goal for change. Much more could be said in developing an account of the instrumental rationality of retributive anger. I have focused on the neglected desire for recognition so as to extend anger's nature and instrumental rationality beyond the narrow (yet important) construal of anger, and its efficacy, as tied to retribution.

Our second desideratum was to provide an account of anger's instrumental rationality. By pursuing a foundational approach to anger, I have challenged the main reasons for taking anger to typically be instrumental irrationality: namely it's modularity and its constitutive link to retribution. I directly challenged the claim that anger is instrumentally irrational in cases of social injustice. The *inefficacy claim* endorsed by Nussbaum, amongst others, amounted to the claim that anger is instrumentally irrational in these cases. I argued that the inefficacy claim is challenged by empirical work as well as everyday moral psychology, and that it should be rejected as a complete, or even paradigmatic, account of anger's relation to instrumental rationality. In so doing, I proposed an account of anger's instrumental rationality that makes the neglected desire for recognition central.

It was highlighted that anger is most effective at bringing about social change when it is perceived by the out-group to be justified, that is, when the angry group is perceived as justifiably angry. This was the appropriateness moderator highlighted in section 3.2. I haven't said much about how to make sense of anger as 'appropriate'. Similarly, in aiming for recognition, anger calls for agents to acknowledge the *reasons* for their anger. I haven't said much about anger as a reason-responsive state. We are in need of an account of anger as normatively rational, or reason-responsive, to make sense of talk of anger being 'appropriate'. In my next chapter, I address this and take steps towards meeting my third desideratum, for the provision of an account of anger's normative rationality.

4

Reasons for Emotions: A Sui Generis Account

*'There is the common assumption that there are only epistemic and
practical reasons...we are talking about the preconceptions of philosophers.
And might it be relevant that they have mostly been male?'*
– Skorupski, in 'The Domain of Reasons'

We typically take emotions to be states for which there are reasons. Victims of sexual harassment have reason to be angry, that your childhood pet has passed away is reason for sorrow, and that you were awarded your doctorate is reason for joy, pride, and perhaps, relief. Similarly, the fact that your friend didn't mean to stand you up last weekend, but actually had a family emergency, is reason not to be angry with him; that your partner belittles and manipulates you is reason to feel harmed by him, and dissatisfied in the relationship, even if you do not.

Our propensity to speak of reasons for emotions suggests that emotions may pick up on facts in the world by virtue of being reason-responsive. Providing an account of anger's normative rationality, then, will involve investigating the mechanics of emotional reason-responsiveness. Recall that a divisive point in the debate on anger, is over whether anger plays a positive, or negative role in rational thinking. I have suggested that 'rational thinking' involves at least two distinct species of rationality: normative rationality, and epistemic rationality. The former concerns reason-responsiveness, while the latter concerns the ability to be knowledge conducive. These two types of rationality are likely to be closely related, after all, anger will plausibly be knowledge conducive in cases where it is responsive to reasons. The relationship between the two, on my account, will come out primarily in my next chapter, where my account of emotional reason-responsiveness will allow anger to play distinctive epistemic roles. For now, I focus on the provision of a sui generis account of emotional reason-responsiveness. This account will allow better sense to be made of the ways in which anger relates to 'rational thinking'.

We often assess our emotions, and those of others, for appropriateness. Is this merely a shallow feature of our folk psychology though? Or does it suggest that emotions are amongst the phenomena of human mental lives that have normative reasons? Normative reasons are reasons that justify. These are reasons that count in favour of having a relevant attitude, but that would also justify that attitude rather than merely explain or motivate it. Explanatory reasons, on the other hand, explain your emotion, though they may not justify it. Normative reasons can be seen as a special kind of explanatory reason, ones that not only make sense of, but also make appropriate or justified, your emotional response. That your friend stood you up last weekend, due to carelessness rather than any family emergency, is a normative reason to be mad at him. That getting mad at your friend is likely to make him never do so again, out of fear of retaliation, for example, might be a motivating or instrumental reason to feign anger, or to try to put oneself in a state of anger, but it does not justify your anger. These types of reasons have been called the 'wrong sort' of reasons, because they are not the proper reasons to which your emotion of anger is itself responsive (D'Arms and Jacobson 2000). At most these reasons justify a belief that one should try to become enraged. The question is whether emotions are reason-responsive in the normative sense outlined[69].

---

[69] I restrict the reasons under discussion to normative reasons, by which I mean justifying or 'the right sort' of reasons. My notion 'appropriateness' is meant to track cases where emotions are responsive to the right sorts of reasons. By appropriate then, I mean what D'Arms and Jacobson (2000) means by 'fittingness', and what

In this chapter I will be concerned with meeting our third desideratum: providing an account of anger's normative rationality. I will argue that emotions are sui generis reason-responsive states. I begin by outlining two main ways of making sense of reason responsiveness; Rationalist Approaches and basic reason-relation approaches. In section 4.1 I argue that the latter approach should be preferred and outline problems with recent exemplars of the former type of approach. Having argued that a basic-reasons approach is in order, I proceed to develop an account of this sort, which is largely lacking in the literature. I begin by supporting my construal of emotions as phenomena capable of reason-responsiveness by appeal to empirical work on attentional dynamics during emotion episodes, in section 4.2. This leaves us in a good position to then characterize the basic reason-relation characteristic of emotions. In section 4.3 I do so, and argue that we have reason to add affective reasons to the widely endorsed dichotomy of epistemic and practical reasons, as a third basic reason-relation with its own irreducible features.

4.1   Emotion and Rationality

Emotions are often cast as being at odds with reason or rational thinking. Such a claim can mean a number of things. There are three senses such a claim can take that are worth teasing apart. First, emotions are claimed to be arational when they are taken to be the sort of state that is not amenable to rational assessment and, therefore, not sensitive to reasons. In addition to this, emotions can be said to be irrational when they fail to be sensitive to reasons, or, when despite responding to reasons they are disruptive to clear thinking. The former involves taking emotions to be capable of reason-responsiveness, but to be failing to respond to reasons at a particular time. The latter involves emotions undermining or biasing one's reasoning, despite being sensitive to reasons for emotions. These two senses of emotion being irrational, failing to reason-responding, and disrupting reasoning, track two different normative concerns, that often both fall under charges of anger undermining rational thinking. What I will follow others in calling 'reason-responsiveness' is a normative relation between an attitude and a fact (Kolodny 2005; McHugh and Way 2018), while what has been called 'rationality' has been taken to involve relations between attitudes being governed by

Danielsson and Olson (2007) call 'content-reasons'. These are notions meant to tease the right sorts of reasons apart from moral and prudential considerations that might count in favour of holding the relevant attitude, but do not speak in favour of the correctness of the attitude.

rational norms or requirements. Rational requirements typically include: coherence amongst ones beliefs, enkrasia (that one intend to Φ when one judges that one ought to Φ), and adherence to rules of inference such as modus ponens (Broome 1999; DeLancey 2002; McHugh and Way 2018)[70].

With this distinction in place, we can appreciate that the question of whether emotions undermine norms of rationality is separate from the question of whether emotions are the sort of attitude that responds to reasons. Fear of failure might compromise your ability to succeed at attaining a life-long dream for example, by causing performance anxiety and inaction that amount to violations of the requirement of enkrasia, or by affecting your ability to adequately apply rules of inference and reason about the likelihood of your success. Similarly, in anger you might form the belief that your friend is a despicable person and unworthy of your time, despite this belief compromising the coherence amongst many of the other beliefs you hold about this friend. If these emotions are arational, then they are not reason-responsive, but they can still cause the described rationality compromising effects, much like arational phenomena such as pains, hunger, and nervous ticks can compromise rational action and deliberation in ways that plausibly amount to violations of rationality requirements.

If your anger and fear are reason-responsive however, they can still be irrational in the sense of interfering with rationality requirements: by biasing downstream inferences and by leading to the violation of rational requirements of coherence and enkrasia for example. Your fear and your anger could also have these rationality violating effects even in cases where these emotions are had for good reason, however. Therefore, although justified emotions might be adequate grounds for beliefs, they can still violate rational requirements. If emotions are reason-responsive attitudes then, their reason-responsiveness is separate from whether they have detrimental (or beneficial, or neither) consequences on the agent's compliance with rational requirements. Rationality, in the norm compliance sense, and reason-responsiveness, come apart.

---

[70] Cohen (2009) lists nine rules of rationality that include: conforming to the laws of deductive reasoning; properly forming theories from inductive cases; making inferences licensed by an accepted factual generalization; performing actions that further the purposes or interests of the agent; choosing the appropriate kinds of ends. See McHugh and Way (2018) for a critique of the rational requirements view of rationality.

How reason-responsiveness and rationality, so construed, relate to one another is the subject of debate. Rationalists explain reason-responsiveness in terms of compliance with rationality norms (Smith 1994; Korsgaard 1996). On such accounts, considerations become reasons only through adherence to the requirements of rationality. It is by virtue of reasoning in conformity with the requirements of rationality that features of the world become reasons for us. Others, take the reason-relation to be basic. This involves taking there to be a basic reason-relation, a normative relation between a fact and an attitude, that doesn't depend on the agent's conforming to any norms of rationality (Raz, 1999; Scanlon, 1998; Kolodny, 2005).

Philosophers of emotion have not engaged extensively with the philosophical literature on reasons, nor have those working on reasons dedicated much of their research agenda to the place of emotions within it. We, therefore, have a gap in the literature on how best to account for emotions as reason-responsive. I will argue that the best way to do so is to construe the reason-relation as basic and independent from deliberative capacities and their compliance with any rational requirements.

I take the main attempts at accounting for emotional reason-responsiveness present in the literature to share Rationalist commitments, in taking reason-responsiveness to depend upon compliance with norms of rationality. I will outline their commitments and highlight a few problems that they face. This will lead me to favour the development of a reasons-basic approach, where emotional reason-responsiveness is a relation between facts and emotions, that hold independently of whether relations between an agent's attitudes comply with norms of rationality.

4.1.1 Agential Disposition Accounts

The most developed accounts of emotional reason-responsiveness in the literature take agential virtues or dispositions to be central (Tappolet 2016; Jones 2003; Goldie 2004b). What these views have in common is that they take reason-responsiveness to rely on properly functioning dispositions to conform to norms of rationality. Agential disposition accounts take emotions to be amongst the agent's reason-*tracking* mechanisms, alongside other mechanisms or 'subsystems' such as the perceptual systems (Jones, 2003). On such accounts

an agent is reason-*responsive* in light of their emotions, so long as she manifests well-functioning reflective self-monitoring habits (Jones, 2003), agential virtues (Tappolet, 2016) or dispositions (Goldie, 2004b). I follow these authors in using these terms interchangeably. The thought is that many sub-systems might track reasons – perceptual systems, emotional systems, perhaps motivational systems – systems we share to large extent with animals, but for reason-*responsiveness* reserved for human agents, there must be well-functioning agential dispositions at play[71]. The crucial disposition invoked on such accounts is the following: so long as the agent is disposed to intervene, and block treatment of their emotion as reason-tracking, when the agent has reason to do so, then the agent's emotions can be seen as properly reason-responsive.

These views are therefore committed to the following claim:

> **Counterfactual Claim:** One is reason-responsive in being epistemically or practically guided by one's emotions, in so far as one would not have treated one's emotion as reason-tracking, had there been reason to believe that the emotion failed to reason-track.

Jones (2003) writes that agents are reason-responsive when:

> the agent's dispositions to reflective self-monitoring are such that she would not rely on that first order sub-system were it reasonable for her to believe that it failed to reason-track … this guidance may remain 'virtual'—that is, revealed in how the agent would behave in various counter-factual circumstances. (195-196)

Similarly, Tappolet (2018) writes that:

> what is required for reason-responsiveness is well-tuned epistemic and practical habits, such that the agent would not act on her emotion had she reason to believe that her emotion mislead her. (499)

Goldie (2004b) writes that reason-responsiveness involves having 'the right habits and dispositions of thought, such that doubts will arise when and only when they should' (251).

The capacity to comply with this counterfactual claim is dispositional. This means that the capacities that confer reason-responsiveness do not involve conscious reflective deliberation but rather a standby sensitivity to when one should engage reflective deliberation, namely when there are reasons to distrust one's emotion. Tappolet (2018b) writes that 'when there

---

[71] Tappolet (2018b) takes reason-tracking to involve merely the detection of reasons, while reason-responding involves being guided, in practical and theoretical reasoning, by reasons.

is no reason to distrust your emotion, you don't need to deliberate to be reason-responsive' (157). Reason-responsiveness will therefore typically be 'unreflective, and not part of conscious deliberation.. (but) rely on our habits and dispositions, at work in the background of our minds, so to speak' (Goldie 2004b: 151).

The counterfactual claim amounts to a claim about dispositional sensitivity to mental-state defeaters. There are two main types of defeaters to beliefs; undercutting defeaters, which give one reason to doubt the truth of the grounds of one's belief, and, rebutting defeaters, which give one reason to hold the negation of the defeated belief, or for holding some proposition that it incompatible with it (Pollock 1986: 38). Believing oneself to be under the influence of drugs is a typical undercutting defeater, while acquiring a belief that contradicts with a belief one holds, is a typical rebutting defeater. Applied to emotions then, we have two types of defeater as well. For an agent to comply with the counterfactual claim, they must be sensitive to when there is reason to believe that the emotion has failed to reason track. This will mean that for one's emotions to be reason-responsive, the agent must be sensitive to when there is reason to believe undercutting or rebutting defeaters are at play. This means being sensitive to whether there is reason to believe that one's emotional system is malfunctioning, perhaps due to the influence of a foreign chemical substance, as well as to whether one has reason to hold a belief that conflicts with the emotion.

This standby sensitivity to comply with the counterfactual claim depends on explicit deliberative capacities however:

> having sensitivity to when reasons are defeated and when they are outweighed requires the capacity to reflect on the status of the deliverances of those mechanisms that purport to latch onto reasons such as perception, emotion and desire, but also the capacity to reflect on reasoning itself (Jones 2003: 190).

The agential dispositions on which reason-responsiveness depend are, therefore, dispositions to engage reflective reasoning when the agent has reason to, where this reflective reasoning is presumably guided by the sorts of norms the rationalist is committed to. Agential Disposition accounts are at the very least committed to one requirement, the counterfactual claim, i.e. that one not trust one's emotions when one believes one has reason not to. Agential Disposition accounts seem committed to the view that emotions are only reason-responsive in so far as agents have dispositions not to violate this requirement, that is, not to treat emotions as reason-responsive when there is reason to doubt that they are. Their reliance on

the capacity for robust reflective reasoning, however, suggests that other rationality requirements are at play in these accounts:

> an agent requires critical reflective ability, dispositions to bring that ability to bear when needed, and dispositions to have the results of such reflection control their behaviour (Jones 2003: 190).

Agential Disposition accounts are therefore still rationalist in a relevant sense, for reason-responsiveness depends on capacities for reflective reasoning governed by rational norms, rather than concerning merely a basic relation between facts and attitudes.

4.1.2 Problems

Both Jones (2003) and Tappolet (2016) take there to be cases of emotion-based akrasia and outlaw emotion that their Agential Disposition account can make sense of as rational. I will argue that Agential Disposition accounts actually face problems in accounting for such cases. First let me briefly outline the sorts of cases I have in mind.

Emotions are considered a main culprit in making agents act against their judgements of how they ought to act (Arpaly 2000; Jones 2003). In anger, we often violate our commitment to civility, fear often prevents us from following a desired plan of action, and in pride or jealousy we can compromise relationships that we are dedicated to the flourishing of. As a common source of akratic action then, emotions seem to be frequently at odds with the enkratic requirement, the requirement that one intend to act according to what one has most reason to do. When Huckleberry Finn fails to act in accordance with his judgement that he should turn his friend Jim in, however, and instead follows a sense of respect and love for Jim, he acts in light of a reason despite acting akratically. Such cases have been called cases of 'rational akrasia' (Arpaly 2000; Tappolet 2016). The thought is that sometimes it is not irrational to violate the enkratic requirement. Many take it to be a condition on a successful account of reason-responsiveness that room for rational akrasia be made (Arpaly 2000; Jones 2003; Tappolet 2016). Any rationalist account is pressed to tell a story for how they can make sense of akrasia as sometimes rational, for on rationalist accounts rationality is a question of conformity to a requirement that akrasia by definition violates.

In 'outlaw emotion' cases (Jaggar 1989), emotions conflict with an agent's wider set of oppressive beliefs[72]. An otherwise content housewife experiences an outlaw emotion when she becomes angry about her confinement to the home, despite endorsing a large set of beliefs about the value of being a housewife, and feeling a range of emotions that cohere with these beliefs, such as pride in fulfilling this role, for example. Recall that Jagger coined the term 'outlaw emotions' to refer to emotions that occur under conditions of oppression, which are 'distinguished by their incompatibility with the dominant perceptions and values' (1989: 166). Much like with rational akrasia, an account of emotional reason-responsiveness should aim to make sense of cases of outlaw emotion as reason-responsive (Arpaly 2000; Jones 2003; Tappolet 2016). Again, rationalist accounts must tell a story for how outlaw emotion cases can be rational, when these seem to be cases characterized by violations of requirements for consistency and coherence.

How do outlaw emotion cases and cases of rational akrasia relate to each other? Cases where one acts on outlaw emotions are cases of rational akrasia, but not all cases of rational akrasia (based on emotion) are outlaw emotion cases. This is because outlaw emotions are specifically one's occurring under conditions of structural oppression. This means that outlaw emotion cases are ones where one's emotion feels 'outlaw' to the agent, as it conflicts with the vast majority of the agent's other beliefs, intentions and emotions. The Huckleberry Finn example outlined above can be seen as one of rational akrasia based on an outlaw emotion. Emotion based akratic actions need not be based on outlaw emotions in the sense intended by feminist philosophers however[73]. That is, they need not be based on emotions that have any important relation to oppression, nor need they be based on emotions that conflict with the majority of one's internalized beliefs. The main reason for teasing these two cases apart is to highlight the practical as well as the epistemic role of emotions as reason-responsive states. Rational akrasia, by name, concerns action, while outlaw emotion cases need not. Indeed, outlaw emotion cases are ones stressed for their epistemic value.

In both cases of rational akrasia, and outlaw emotion, there are arguably rebutting defeaters at play. One plausibly has reason to believe a proposition that conflicts with the emotion in

---

[72] See chapter 1, section 1.2.1. In my next chapter, on the role of emotions in justifying beliefs, I will give outlaw emotions much fuller treatment. For now, I mean only to highlight the difficulties Agential Disposition accounts face in trying to make sense of such cases as cases where the outlaw emotion is reason-responsive.

[73] Arpaly (2000) makes use of a number of examples of emotion-based akratic action that are not outlaw emotions in the sense intended by feminist philosophers. That is, they are not 'outlaw' by virtue of conflicting with an agent's wider set of *oppressive* beliefs.

both cases. In rational akrasia, one has a practical belief about what one ought to do, in Huck's case that he should hand Jim in, which conflicts with one's emotion of love for Jim, and in outlaw emotion cases, one endorses a normative belief about the value of being a housewife, for example, which conflicts with the outlaw emotion of anger about being a housewife. In both cases the agent has at least some reason to believe that their emotion is failing to reason track, given the presence of these conflicting beliefs which seem to be rebutting defeaters to one's emotions. That is, they give the agent reason to doubt that one's emotions are properly tracking reasons.

Both Jones (2003) and Tappolet (2016) take the Agential Disposition account of emotional reason-responsiveness to be able to account for cases of rational akrasia and outlaw emotion, however. On rational akrasia they are explicit:

> The functioning of such sub-systems does not stop being expressive of our commitment to rational guidance just because there is now an opposing all-things-considered judgment. In some cases that all-things-considered judgment may be such that the agent would distrust it, if her self-monitoring capacities were functioning as they should. Thus, the regulated sub-system can be more expressive of the agent's commitment to rational guidance than the all-things-considered judgment: the incontinent action can display the agent's commitment to rational guidance more fully than does the continent action. (Jones 2003: 196)

> It follows from this account that an akratic agent, who is motivated by her emotion to act against her practical judgement as to what to do, manifests reason-responsiveness. This is so when in spite of her practical judgement, it would not be reasonable for the agent to believe that her emotion fails to track her practical reasons. (Tappolet 2018: 449)

An emotion's reason-responsiveness is, therefore, not undermined by its conflicting with an all things considered judgement of how one ought to act, in cases of rational akrasia. This is because, according to Tappolet at least, such conflicting judgements are not *themselves* reasons for the agent to distrust one's emotion. Jones (2003) likely endorses this as well, as she says that akratic actions are rational when:

> (1) the action is produced by a sub-system that reason-tracks because the agent reason-responded, and (2) the agent would have distrusted her all-things-considered judgment were her self-monitoring dispositions operating as they should. (196)

For (1) to hold, the agent must not have had reason to believe the emotion was failing to reason-track, given the account's commitment to the counterfactual claim. Agential

Disposition accounts would presumably make sense of outlaw emotion cases in an analogous way, that is, by taking the conflicting judgement to not constitute a reason to doubt one's outlaw emotion is reason-tracking.

I think Agential Disposition accounts are too quick to assume success in accounting for cases of rational akrasia and outlaw emotion. Firstly, consider Jones' (2003) two requirements for the possibility of emotion-based rational akrasia. The first involves the agent manifesting well-functioning self-monitoring dispositions, this is what allows the emotion 'subsystem' to 'reason-respond', while the second suggests that the agent actually has malfunctioning self-monitoring dispositions. Well-functioning agential dispositions would not have allowed the agent to endorse the all-things-considered judgement with which the emotion conflicts. Therefore, if the agent holds such an all-things-considered judgement, they are not manifesting well-functioning self-monitoring dispositions. Cases of rational akrasia, according to this view then, seem to involve the unattractive result of an agent both having well-functioning agential dispositions (that grant the emotion reason-responsiveness), while manifesting malfunctioning self-monitoring dispositions (in holding the all-things-considered judgement), simultaneously. In so far as these dispositions are agential, it seems that one either has well-functioning dispositions in a given circumstance, or not.

But there is a further worry, for if we assume that one is manifesting well-functioning dispositions with respect to one's emotional reason-tracking system, and therefore that one's emotions are reason-responsive in cases of rational akrasia, as Tappolet and Jones claim, then it isn't clear why the conflicting all-things-considered judgement wouldn't count as a reason to doubt the emotion is properly reason-tracking. It seems that if one is operating with well-tuned agential dispositions, then cases of emotional akrasia would be ones that should be picked up on by the agent's reflective self-monitoring systems as requiring further deliberation. This is because, in such cases, one acts, or intends to act, against what one takes oneself to have most reason to do. One, therefore, has a judgement that conflicts with one's intention. This judgement is likely a rebutting defeater against one's emotion and therefore prone to being picked up by well-functioning dispositions for norm governed reasoning.

If no deliberation has been triggered in such cases then one's reflective self-monitoring has failed to pick up on the conflict between one's emotion-based intention and one's judgement of what one ought to do, and one's emotions are therefore not reason-responsive because

one is not manifesting well-tuned self-monitoring habits. If, on the other hand, deliberation is triggered, due to well-functioning agential dispositions, then, either the akratic situation is dissolved, or, one is irrationally akratic. If deliberation is triggered, then, either one's initial judgement of what one ought to do is abandoned (or revised), or it is maintained. If the result of deliberation is such that the agent comes to endorse the once akratic action and abandon one's initial judgement of what one ought to do, then, one is no longer akratic as one is now inclined to act in light of what one believes one has most reason to do. Therefore, one is either rational or akratic on such accounts, there seems to be little room to account for rational akrasia.

If we want to grant that there are cases of rational akrasia, then it seems that what makes akratic action rational cannot be its dependence on the agent's dispositions for reflective-self monitoring, and the capacity to intervene when they have reason to doubt the emotion is getting things right. After all, as we have seen, akratic cases are ones that well-functioning agential dispositions plausibly should tag as in need of revision. Revision would dissolve such cases, therefore, akratic action involves failures of such agential dispositions. In so far as we continue to believe some akratic action rational, then, what makes such actions rational must be something other than a dependence on such agential dispositions for norm governed reasoning[74]. The fact that there is a reason for the emotional attitude itself seems to be the answer.

Similarly, in the epistemic, outlaw emotion, cases that feminists have highlighted, the outlaw emotion seems to be reason-responsive despite conflicting with the agent's wider set of beliefs. Can Agential Disposition accounts make sense of this, however? On the agential dispositions account, outlaw emotions would presumably be flagged as problematic for violating coherence norms and conflicting with beliefs (as well other attitudes including emotions and intentions). The beliefs with which the outlaw emotion conflicts are presumably reasons the agent has for thinking the outlaw emotion is failing to reason-track. If this is so, then, if one is displaying well-tuned agential dispositions, one's stand-by monitoring processes would intervene and trigger deliberation in outlaw emotion cases. This means that outlaw emotion cases are not reason-responsive on Agential Disposition accounts, as they violate the counterfactual claim.

---

[74] Arpaly (2018) makes a similar point in response to Tappolet's (2016) account.

Outlaw emotion cases, on these accounts, would arguably only be reason-responsive in cases where the deliberation triggered resulted in an endorsement of the outlaw emotion. This risks dissolving outlaw emotion cases, however, as these are cases where the agent often lacks the conceptual resources to arrive at a deliberative conclusion that would make sense of the emotion, given their internalization of oppressive ideology[75]. Agential disposition accounts would, therefore, at the very least, limit the cases in which outlaw emotions can be said to be reason-responsive. Any case where the agent cannot make deliberative sense of the outlaw emotion would be one in which the outlaw emotion is not reason-responsive. This is an undesirable outcome. To make sense of outlaw emotion cases as rational, the fact that there is a fact that stands in support of the emotional attitude itself must be invoked.

Agential Disposition accounts run into problems accounting for cases of outlaw emotion and rational akrasia. Proponents of such views must say more on how they intend they intend to account for such cases[76]. I take sufficient reason to have been given to warrant exploration of an alternative account of emotional reason-responsiveness, specifically one that makes sense of reason-responsiveness as basic. We saw that to make sense of cases of outlaw emotion and rational akrasia, it is intuitive to appeal to such a basic relation. I have not claimed that there are no moves available to the Agential Disposition account to make sense of such cases, but I take the difficulties they encounter in doing so to suggest a departure from Agential Disposition accounts, towards basic-reason accounts, to be a fruitful direction to take.

Although some philosophers of emotion seem to speak of the reason-relation as basic (Deonna and Teroni 2012) they do not provide explicit endorsements of this, or detailed accounts of what this basic relation involves. In the rest of this chapter, I attempt to do just that. I begin by motivating the initial assumption that emotions are the sort of thing that can be had for reasons, where reasons are facts that bear a basic reason-relation to an attitude.

---

[75] See section 1.2.1 in Chapter 1. I will return to this point in the next chapter where I focus on the epistemic role of outlaw emotions in providing reasons for beliefs.

[76] Tappolet (2018b) provides a response to the sort of objection I have raise here. Tappolet takes the conflicting judgements in cases of outlaw emotion and rational akrasia to not be proper reasons to doubt one's emotions are reason-tracking. This is because, Tappolet claims, these conflicting judgements are not direct evidence for the malfunctioning of emotions, they are therefore not standard undercutting defeaters (such as being drunk or under hypnosis), nor are they rebutting defeaters, because they are not *justified*. As Tappolet's response hinges on epistemological commitments regarding the nature of justification and defeaters, I deal with it in chapter 5, where I focus on the epistemic role of emotions. There I argue that Tappolet's moves to deny that outlaw emotion cases involve defeaters fail. For now, I will say that it isn't clear that standard undercutting defeaters are the only defeaters at play in such cases, nor that the conflicting judgements typical of outlaw emotion and rational akrasia cases are unjustified.

4.2 Emotions as Reason-Responsive

For emotions to be reason-responsive, they must be cognitive phenomena, they must have rationally evaluable content (Deonna and Teroni 2012; Tappolet 2016). That emotions are intentional states is now the dominant view in the literature, however, so this requirement only immediately excludes classic feeling emotion theories (Deonna and Teroni, 2012). Mere intentionality is not sufficient for reason-responsiveness however. Some intentional states are not rationally evaluable, perceptions are a paradigmatic example (Brady 2013). Perceptions are taken to be truth-apt but not rationally evaluable, that is, we may have causal reasons for perceptions but not normative ones. We can be praised or condemned for our emotions, but typically not for our perceptions. One is not considered irrational for misperceiving, while one is often considered irrational for experiencing an emotion for which one lacks reasons.

This distinction, between emotions and perceptions, is obscured on Agential Disposition accounts. On such accounts, both emotional systems and perceptual systems are reason-tracking, that is, they both passively register certain features of the world. Reason-responding, as opposed to mere reason-tracking, depends on treating reason-tracking sub-systems in a manner that manifests well-functioning dispositions for reflective deliberation. On these accounts then, emotions and perceptions are taken to bear similar relations to the reasons they track. The intuition that emotions are had for reasons, while perceptions are not, is not accounted for on these accounts, at least in the first instance. The observation that emotions seem to admit of justification themselves, while perceptions typically do not, should alert us to the possibility that emotions are not mere reason-trackers.

On a reasons-basic approach, an attitude is either the sort of kind that is had for reasons, or it is not. On such accounts, whether one is reason-responsive in having the relevant attitude depends on whether there is reason, typically construed as a fact, for it. A reasons-basic approach allows us to make initial sense of the difference between perceptions and emotions that Agential Disposition accounts obscure. On a reasons-basic approach, emotions are the sort of attitude for which we have reasons, that is, the sort of attitude that bears a reason-relation to facts, while perceptions are not. Perceptions might be said to 'track' reasons, but they (typically) are not taken to bear reason-relations to facts[77].

---

[77] See Siegel (2017) for a recent departure from this orthodox view.

Although perceptions are perspectival, that is, relative to various situational constraints, and also more cognitively penetrable than classically had been thought, the label of reason-tracker doesn't seem ill-fitting. Emotions, on the other hand, are taken to involve an evaluation, the taking of a stance, towards which facts count in favour or against. This is where emotional reason-responsiveness bears similarity to the epistemic and practical cases. Emotions, much like beliefs or intentions to act, involve taking a stance on something. In the epistemic case, one takes the stance that a certain proposition is true, while in the practical cases one takes a stance on what to do. Both involve bearing a particular type of relation to a proposition, one that goes over and above the mere registering of the contents of that proposition. Emotions may be much the same. In having an emotion, one takes a stance on an object or state of affairs, specifically an affective, rather than epistemic, or practical stance. Facts can count in favour, or against, such stances.

While in the epistemic and practical cases, theoretical and practical reasoning are conscious and often deliberate phenomena, in the case of emotions the evaluative stances taken are thought to be underlied by unconscious appraisals (Scherer 2005). Emotions can, therefore, be thought of as the output of subconscious processing on the evaluative significance of happenings in the world. We will shortly see that these appraisals are dynamic, that is, emotions involve reappraising situations as they occur (Scherer 2005; Deonna and Teroni, 2012). We will see that emotional appraisals function in a manner that is not encapsulated from background information and the agent's wider concerns. Empirical work on attentional dynamics during emotions will lend support to such a view.

It is this taking of an evaluative stance, in emotions, that makes them the sort of kind for which there are reasons. Recall that in chapter 2 I stated my agnosticism concerning whether the evaluative component of emotions is part of the content of emotions, or a feature of the type of attitude that they are. I remained, however, committed to emotions having evaluative *phenomenal* content, whereby emotions involve feeling their objects to instantiate relevant evaluative properties. The phenomenology of emotions often involves attentional focus on the features of the object of one's emotion, that are felt to merit the relevant evaluation.

When angry at your boss for making you alone work late, for the third day in a row, your attention focusses on your boss, whether perceptually or imaginatively. Your eyes follow

your boss around the office, attentive to their moves. Once home, your mind is busy replaying the offense you have suffered, re-hearing the tone of your boss' voice and their manner of speech. When in anger, your attention is focused on the offence you have suffered and the person or social object responsible for it. Similarly, when fearful of a dog, your visual attention latches on to it and its sharp teeth. It is hard for you to attend to anything other than the object of your fear. Emotions such as fear and anger often interrupt your behaviour and consume the vast majority of your attentional capacities, such that concentrating on anything else becomes difficult. Indeed, the propensity of emotions to distract us from daily endeavors is a common charge against their rationality: 'I can't concentrate, I'm too angry!'. That this is so is a clear indication that emotions and attention are intimately linked.

But it is also this engagement of attention that allows emotions to be reason-responsive. Emotions involve attending to considerations that count in favour of them. Attention is key to differentiating emotional from non-emotional phenomena. One can evaluate objects and states of affair without experiencing emotions. One can believe that one has been offended, or that a situation merits sorrow, without experiencing the emotions of anger or sadness. Note that this is a claim about instances of applications of evaluative concepts, rather than a claim about the capacity to apply evaluative concepts. The former is possible without experiencing an emotion while the latter might not be. The difference between taking something to be dangerous, such as having a belief to this effect, and feeling fear, is that in one case you feel a certain way and in the other you do not. Emotions have a distinctive phenomenology, while evaluative beliefs typically do not. A crucial component of emotional phenomenology is attentional. Emotions consume our attention, distracting us from other concerns, they can saturate the entirety our attentional capacities. This is part of the phenomenological difference between having an evaluative belief that something is dangerous and feeling an emotion of fear towards the same object. Evaluative beliefs do not engage attention as strongly, or in the same ways, as emotions do.

The strong link between emotion and attention has been noted by psychologists and philosophers (de Sousa 1987; Faucher and Tappolet 2002; Frijda 2005; Brady 2013), yet the significance of this link has not garnered as much philosophical attention as it deserves[78]. I argue that an important way in which attention operates during emotions is that one's attention typically becomes directed toward considerations that appear to count in favor of

---

[78] Exceptions include Faucher and Tappolet (2002) and Brady (2013).

the emotion[79]. Part of an emotional experience of anger then, is to have one's attention insistently drawn to reasons for offence.

Having an emotion E, such as anger directed at object O, is not just a matter of seeing something wrong about O. You can think something is wrong or unjust without getting angry about it. Emotions involve a focusing of attention such that to have emotion E involves the agent's attention being drawn to considerations that present themselves in favour of E[80]. These considerations are reasons for which we have the emotion E.

The idea that attention plays a key role in emotional reason-responsiveness is not novel. De Sousa (1987: 196), for example, writes that 'paying attention to certain things is a source of reasons', and Brady (2013: 160) takes the 'capture and consumption' of attention during emotional experiences to allow us to discover reasons for them. To cash out, and further specify, the claim that emotions involve a direction of attention towards the reasons that support them, we need to have a better understanding of the relationship between attention and emotion. Relatively little has been written on the relationship between the two in philosophy, notable exceptions include de Sousa (1987), Faucher and Tappolet (2002), and Brady (2013). On the other hand, there is a vast literature within empirical psychology and neuroscience on this topic. A brief examination of some of the main findings in these disciplines will lend support to my construal of emotions as reason-responsive states.

4.3   Emotion and Attention

To talk about the relationship between two phenomena, emotion and attention, we better get a bit clearer on what we are referring to in each instance. By emotion, I will continue to

---

[79] Note that this does not imply that the agent must conceive of reasons as reasons, nor need she conceive of an offence she has suffered as an offence, for her anger to keep the offence salient in her mind.

[80] It has been noted that although negative emotions narrow one's attentional focus, positive emotions actually do the opposite, and widen one's attentional scope (Fredrickson and Branigan 2005). This observation does not pose a challenge to my proposal, however. In the first instance, as I am interested in anger, this experimental evidence does not challenge my proposal's applicability to our target emotion. We have reason, however, to doubt that the experimental work gives us reason to treat positive and negative emotions differentially. Firstly, the authors of the widely cited experimental paper on this, themselves admit that the evidence for differential attention in positive versus negative emotions to be largely inconclusive. Secondly, as Brady (2014) highlights, many of the experimental results pertain to the downstream effects emotions have on attention, that is, to how emotions affect our attention to objects in our environment other than the objects that trigger the emotion. The experiments, therefore, do not focus on how attention to the objects of emotions differs between emotion types. See Brady (2014) for a thorough account of why this experimental evidence does not pose a problem to the proposed relationship between attention and emotion.

mean short-lived occurant episodic states. I am, therefore, not talking about moods, dispositional affective states, or emotional character traits, although all of these phenomena may be related to attention in distinct and interesting ways, I focus on paradigmatic short lived, phenomenologically salient emotion episodes that have intentional objects.

Attention, not unlike common uses of the word emotion, is somewhat of an umbrella term for a number of related phenomena. Unlike with emotion however, where I am focusing on only one specific subset of emotional phenomena, multiple forms of attention will be relevant to our concerns[81].

One axis along which attention is meant to vary, relates to the source of attentional processing. Two distinct causes of attentional changes are typically postulated in the empirical literature: exogenous attention and endogenous attention. The former is characterized as 'bottom-up', 'stimulus-driven' attention, while the latter is 'top-down' or 'goal-driven' attention (Fazekas and Nanay 2018). The difference has often been cashed out as whether attention is mobilized by external or internal factors[82], that is, by external stimuli or by the agent, in accordance with their goals. Bottom-up or stimulus-driven attention is paradigmatic of the emotional realm, the empirical literature takes this to be because emotional stimuli are ones with hard-wired evolutionary significance. This means that responding emotionally to a given stimulus involves having one's attention automatically drawn to the object of the emotion. Bottom-up attention is typically taken to be recruited quickly, automatically, and independently to voluntary control (Faucher and Tappolet 2002; Pessoa 2010; Carretié 2014).

It is important to note that 'automatic' is not necessarily the opposite of 'voluntary'. Both terms have been used quite differently across the brain and behavioural sciences. For our purposes we will use 'voluntary' to mean attention that is consciously willed by the subject. As emotions are not states that we can will ourselves into, much of the attentional processing

---

[81] Throughout I will be making use of similar distinctions to those made in Faucher and Tappolet (2002), whose work represents one of the most in-depth considerations of the relation between emotion and attention in the philosophical literature. I will, however, be expanding upon their work in two important ways. First, by bringing more recent empirical work to bear on the relationship between emotion and attention, particularly work on the temporal profile of attentional processes during emotional episodes which they did not broach, and second, in applying this evidence to the issue of emotional reason-responsiveness, which was not their focus.

[82] This is distinct to whether the object of one's emotion is 'internal', such as when one becomes saddened by a memory for example, or 'external', such as when you are saddened by a current state of affairs. The empirical work discussed refers to the latter sort of cases. I take the main findings to be in principle applicable to emotions bearing 'internal' objects of this sort. It will not be my concern to argue for this here, however.

occurring during emotional episodes will be involuntary. This isn't to say that these attentional processes will all be *automatic*, however. For the empirical work we are concerned with, 'automatic' is best construed as meaning processing that is independent to significant top-down influences (Pessoa 2010)[83]. When processing is automatic, then, it will be independent to the agent's goals, intentions and character. Stimulus-driven attention is therefore automatic and involuntary in these senses, while endogenous attention is not automatic, and can be recruited both voluntarily and involuntarily.

Whereas stimulus driven attention is paradigmatically involuntary, the reverse is not straightforwardly true of top-down attention. Although voluntary attention will typically be top-down, given that volitional and explicit attention arises endogenously and correlates with activity in higher-level neural processing areas, top-down attention is not typically voluntarily. When attention is recruited, or maintained, by higher level processing areas sensitive to the agent's goals, memory and character, this will likely be endogenous attention, whether or not an agent has explicitly willed this (Pessoa 2010). This brings me to a final distinction, that between the shifting of, and the maintenance of, attention. Both can be undergone voluntarily or involuntarily.

I propose that we take the bottom-up/top-down distinction to be a difference in the source of attentional processing, while the voluntary/involuntary distinction involves a difference in the type of attention at play, and shifting/maintenance reflect different phases of attention. These three dimensions operate largely independently, in that, sources, types, and phases of attention do not come in mutually exclusive triads. That being said, voluntary attention is necessarily top-down and bottom-up attention is necessarily involuntary.

Of particular relevance to us, in our discussion of how emotions are reason-responsive, is one type of attention: involuntary attention. We will be concerned with both sources as well as both phases of involuntary attention. While we can, of course, voluntarily attend to, or search for, reasons for our emotions, consciously and explicitly, and this might be an important way of gaining evaluative insight[84], emotions themselves typically arise involuntarily, such that in so far as we are interested in emotions themselves as reason-responsive, our focus should be on involuntary attention.

---

[83] 'Automaticity', as it has been used in the attention literature, can be seen as akin to the Fodorian concept of informational encapsulation.
[84] See Brady (2013) who takes this to be the main epistemic role emotions play.

4.3.1 Stimulus-driven attention

Empirical evidence supports our intuitive view that objects of particular significance to the organism involuntarily, automatically, and passively recruit attention. It is typically assumed in the empirical literature that emotional stimuli are intrinsically salient because of their evolutionary significance. Stimuli can, alternatively, be conditioned over child-rearing to recruit attention automatically as well. Stroop tasks and pop-out tasks provide evidence for emotionally salient stimuli drawing attention automatically. In the original Stroop task, participants had to name the colour in which different words were written as quickly as possible (Stroop 1935). Some of the words corresponded to the colour they were written in; the word 'PINK', written in pink, is an example of a congruent stimuli because the colour of the word matches the semantic content. Incongruent stimuli were those where the colour of the word and the semantic content did not match; the word 'PINK' written in green for example. Results showed participants to take longer to name the ink colour of the words when faced with incongruent stimuli compared to congruent stimuli. This has been interpreted as indicating that semantic content competes with visual information for attention, resulting in a slower response time (Mathews and MacLeod 1985).

An emotional analogue of this experiment has been devised where subjects are asked to identify the ink colour of words that are either emotionally neutral (e.g., 'apple') or emotionally salient (e.g. 'death') (Mathews and MacLeod 1985). Responses for emotional words showed higher response times. This supports the observation that emotionally significant stimuli compete for attentional resources by drawing more attention than neutral stimuli. In pop-out tasks, subjects are exposed to images containing various objects and asked to indicate whether all of the objects are the same (e.g. multiple daffodils) or whether the image contains an 'intruder' (e.g. one mushroom amongst many daffodils). In some trials the intruder was emotionally relevant (e.g. snake). Faster response times were found when the intruder was emotionally relevant compared to when the intruder was neutral or to when there was no intruder. This has been interpreted as evidence that emotionally-relevant stimuli are quicker to draw attention than neutral stimuli (Carretié 2014).

These experiments do not specifically isolate stimulus-driven attention, as the subjects are instructed to attend to relevant stimuli, which involves top-down involvement. There are other paradigms that select for bottom-up effects that confirm the evidence above, however.

Concurrent but distinct target distractor (CDTD) paradigm tasks involve the endogenous attention of subjects being directed at a particular task, set by the experimenter. Usually this involves judging whether two images are the same or two lines are the same length. While the participants are busy working on this task, 'distractors' appear on the screen that are emotional on some trials, and non-emotional on others. Distractors are often faces that are either neutral or emotional. Studies show greater attention is drawn to emotional than non-emotional distractors. This has been confirmed by behavioural measures, such as decreased response times, when the distractor is emotional, as well as by tracking eye movements, where emotional distractors are looked at for longer (Carrieté 2014). This suggests that emotional stimuli are more distracting than non-emotional ones, due to their automatic recruitment and consumption of attention.

The temporal profile of attention during emotional experiences is important to our concerns. Recent findings using EEG allow neural activity to be investigated with excellent temporal acuity. Findings indicate that bottom-up attention is quickly mobilized upon the detection of emotionally relevant stimuli (around 100ms after detection of stimulus). A pre-attentive evaluation is thought to occur slightly before this, when the discrimination between emotionally relevant as opposed to irrelevant stimuli is thought be made. In recruiting attention, various reorientation networks are activated that automatically reorient the individual so as to attend to the object more fully. This involves ocular, neck, head, and body movements, as well as the decreasing of processing of other environmental factors. This reorientation makes sensory amplification possible. Sensory amplification involves preferential sensory processing of the target stimulus, allowing an enhanced representation of it to be built up. Sensory amplification commences as soon as reorientation occurs. The pre-attentive evaluation is thought to take place in limbic structures such as the amygdala and insula, which in turn connect to the circuits that orchestrate reorientation as well as the visual cortex where activity is modulated to potentiate information gathered from the relevant stimuli (Carrieté 2014).

4.3.2   Endogenous Attention

Around 300ms after stimulus onset, endogenous attention kicks in, which involves attention processing from higher brain regions (Carrieté 2014). In decision making paradigms there is typically a peak in ERPs (event-related-potentials, the data recorded via EEG) at around 300ms after stimulus presentation. This peak is called the P300 and is thought to reflect endogenous attention, as it is more correlated with the individual's response to the object, as opposed to stimulus detection. The P300 is thought to reflect stimulus evaluation or categorization (Clayson and Larson 2013). It is, therefore, perhaps unsurprising that the P300 generated by emotional stimuli is stronger than that of neutral stimuli (Hajcak et al. 2010). The emotional P300 is modulated by conflicting emotional stimuli, which suggests that it correlates with processing involved in emotional conflict resolution (Clayson and Larson 2013). Indeed, the same higher-level brain areas involved in top-down attention, the ACC and vmPFC, have been implicated in emotional conflict resolution within the same time frame as the P300 (Etkin et al. 2006).

There is, therefore, evidence that areas responsible for categorization and conflict resolution are recruited in endogenous emotional attention. As this activity has been tied to the resolution of emotional conflict, it provides evidence for the existence of self-regulatory processing within or during emotional episodes (Carretié 2014). All attentional components discussed commence within a 300ms window from the onset of the stimulus. This speaks in favour of taking these attentional processes to be components of emotional episodes themselves. This picture is supported by appraisal theories where re-appraisal mechanisms are taken to be part of the emotional episodes (Scherer 2005). On such accounts an emotion is a dynamic episode where the initial appraisal can be modified following the gathering of further, perhaps disconfirming evidence[85]. Fig. 1 illustrates the temporal profile of attention during a typical emotion.

---

[85] For a review of emotion regulation mechanisms see Hajcak et al. (2010)
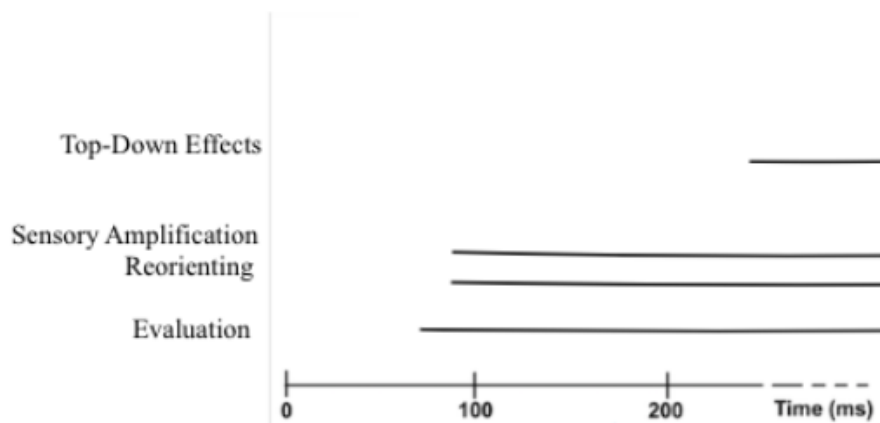
Figure 2 Temporal profile of attention during emotion (modified from Carretié 2014)

What this suggests, is that it seems to be part of emotional episodes for attention to firstly be triggered in a bottom-up fashion by the stimulus, where this attention serves to gather further information about the stimulus, and then for top-down attention to be recruited, which essentially 'checks' whether the gathered information matches the initial emotional appraisal or categorization. The features of the stimulus that are being subjected to sensory amplification are ones that presumably support one emotional categorization over another. Which features of the stimulus are preferentially processed, or amplified, is under the influence of top-down attentional demands once endogenous attention kicks in (Fazekas and Nanay 2018). There is, therefore, an influence of endogenous attention on sensory amplification. When there is conflicting emotional information present in the stimulus, ACC has been observed to be triggered around 300ms after stimulus presentation (Clayson and Larson 2013). This suggests that, although emotional stimuli recruit bottom-up attention automatically, soon after, endogenous attention is recruited to regulate, or 'check', whether the stimulus merits the initial emotional categorization given by the automatic appraisal, as well as to resolve any categorization conflicts that may arise. This may involve regulating which sensory inputs are amplified. These regulatory functions are involuntary but not automatic in the sense described above: they are not independent from top-down input from higher brain regions, seeing as they are regulated by regions responsible for categorization. This means that emotional episodes seem to involve attentional processing that is not encapsulated from the individual's intentions, goals, memory and character.

A closer look at attentional processing during emotional episodes, therefore, suggests that emotions involve endogenous attention that mediates pre-verbal and involuntary categorization and evaluation of stimuli. This endogenous attention is mediated by higher level brain regions that are not likely to be encapsulated from the agent's wider aims, memory, character and goals. It is indeed likely that these factors feed into and influence the categorization of emotional stimuli (Pessoa, 2010).

Emotions, therefore, do not seem to merely track information in a manner encapsulated from the agents wider aims, memory and character[86]. Emotions seem to be the sort of attitude that facts count in favour of as they seem to include re-appraisal mechanisms that evaluative whether the initial categorization of an emotional stimuli is correct.

4.3.3 Taking Stock

Recall your anger at your boss for singling you out to work late again. Your attention became directed at features of this offense, as well as towards the tone of your boss' voice and their manner. In anger, your attention became directed towards the target of your emotion, your boss, and shifted away from other concerns. You began to automatically and selectively amplify sensory information, that is, your representation of the boss' words, voice, and manner were enhanced. In having your emotion drawn to considerations that count in favour of your anger at your boss, your endogenous attention is engaged in reappraisal, checking whether the features that have been amplified merit the relevant evaluation. These reappraisal mechanisms occur in higher-brain regions that are not encapsulated from your wider interests, desires, or knowledge. The appraisal involved in your anger is, therefore, likely to be confirmed, or even intensified, once your memory of past events is brought to bear on the initial appraisal. After all, it is the *third* day in a row that your boss has made you work late. The fact that working late will make you miss a dinner with a friend, which you had been looking forward to, will also contribute to the strength of emotion felt appropriate to the state of affairs.

I have supplemented the intuitive idea that emotions are reason-responsive attitudes with a picture of emotional episodes, and their underlying processing, that supports this stance.

---

[86] Indeed, this fits with what we have gathered from chapter 2 where the modularity of anger was challenged.

Emotions then, like other characteristically reason-responsive attitudes, such as beliefs and intentions to act, are subject to reappraisal and revision on the basis of further relevant information, including the agent's background knowledge and aims. Emotional reason-responsiveness seems to be more than a feature of human language. The evidence just surveyed supports an account of emotions as dynamic reason-responsive attitudes.

Attentional dynamics during emotional episodes seem to mediate at least some forms of unconscious 'counting' of the features initially highlighted, in favour of certain evaluative categorizations. Indeed, the attentional processing involved in emotions selectively amplify the representation of relevant features of the object, and 'check' whether the initial appraisal is warranted in light of these enhanced representations. These selectively amplified features of the object of one's emotion are reasons that count in favour or against the initial appraisal.

The role I take attention to play during emotional episodes fits well with the few treatments of the relationship between emotion and attention in the philosophical literature. Brady (2014) specifically writes that

> attentional persistence can enable or facilitate an enhanced representation of potentially significant objects and events: by keeping our attention fixed on some object or event, emotions can enable us to get a better grasp of our evaluative situation, by allowing us to discover reasons or considerations that bear on the accuracy of our initial, reflexive, and automatic emotional appraisal (54).[87]

Emotions therefore do not merely highlight facts, as Agential Disposition accounts imply in grouping emotions along with perceptions, as reason-trackers. Emotions seem to involve internal attentional processes that make them reason-responsive attitudes, as opposed to mere reason-tracking attitudes that depend on dispositional compliance with norm governed reasoning to grant them reason-responsiveness.

---

[87] I take my view to be quite in line with Brady's. Brady (2013) takes emotions to be temporally extended phenomena that involve attention in ways that allow them to play crucial epistemic roles. As do I. But Brady seems committed to the view that the consumption of attention involved in emotions plays this role by motivating us, as agents, to assess and search for reasons for the emotion. I don't deny this, but I take the crucial role of attentional dynamics during emotional episodes to be that they underlie unconscious reappraisal. That is, emotions themselves involve 'checking' reasons for them, rather than, or in addition to often, prompting agents to search for reasons for them. It is not entirely clear to me whether Brady would reject the picture I propose. One reason for thinking he might, is that Brady (2013) takes emotions to raise justificatory concerns about themselves. On my account emotions do not, this is because I don't take this to match the everyday phenomenology of emotion experience (as I will argue in chapter 5), but also because it seems to be part of emotions *themselves* to include reappraisal mechanisms that 'check' whether the emotion involves the correct appraisal. This supports a view emotional phenomenology that I favour, where emotions typically, or at least often, involve feeling that they are appropriate, in virtue of certain aspects of their objects.

Emotions share similarities with other attitudes that are thought to bear basic reason-relations to the facts that support them, namely beliefs and intentions. While we have seen that there are good grounds for thinking of emotions as genuinely reason-responsive, their reason-responsive character is different, in important ways, from that of belief and action. The most notable difference being that theoretical and practical reasoning, which are common ways of responding to epistemic and practical reasons, are typically undergone under the command of the will, and are often explicitly engaged in. Emotional reason-responsiveness, on the other hand, typically involves unconscious processes, including appraisal mechanisms, the steps of which are not transparent to the agent nor under the command of their will. Is this the only relevant distinction between emotional reason-responsiveness and epistemic and practical reason-responsiveness? We have reason to think not. In the next section I argue that the emotional reason-relation should be construed as sui generis, that is, as irreducible to either practical or epistemic reason-relations. This will lead us to the endorsement of a trichotomy, rather than a dichotomy, of basic reason-relations.

4.4   A Trichotomy of Reasons

Amongst those that support construing the reason-relation as basic, there is widespread endorsement of two types of basic reason-relation: reasons for belief and reasons for action (Raz 1999; Skorupski 2010; Maguire 2018). Amongst those who propose reason-relations as basic, it is common to construe reasons as facts that count in favour of certain attitudes (Raz 2011; Scanlon 2014). I follow them in doing so[88]. The reason-relation is basic in that the relation between facts and attitudes does not depend on the agent's conformance with norms of reasoning. On reason-basic accounts, reason-relations are properties of facts that stand independently to the agent's reasoning practices. In the literature on the emotions, we find widespread endorsement of emotions being reason-responsive (Greenspan 1988; Deonna and Teroni 2012b), but this stance has typically only been systematically developed in the form of the Agential Disposition accounts I dealt with earlier. In the literature concerning

---

[88] Emotions are justified or appropriate when there are normative reasons, that is, the right sort of reasons in their favour. By construing reasons as facts, this means that when emotions are responsive to *apparent* reasons they are not appropriate, though they might still be intelligible or warranted. Fear of a shark is generally justified, for example, while fear of a fake shark is only warranted. Similarly, anger at your friend is justified if your friend has offended you, but it is still warranted or intelligible if you have apparent reason to think they have offended you (though in fact they have not).

reasons, emotions are rarely ever the focus, and general endorsements of emotions being reason-responsive states are harder to come by[89].

This leaves a gap in the literature on the nature of the normative support of emotions. It also leaves alive the possibility that emotions might be reason-responsive in much the same ways as other reason-responsive states, such as beliefs or intentions. The remainder of this chapter is concerned with arguing against this possibility, and carving out a distinctive reason-relation for the emotional realm. I call the reasons that justify emotions 'affective reasons'.

Like reasons for action and belief, reasons for emotions are normative, they make the relevant evaluation appropriate to the object in question. The central difference between epistemic, practical and affective reasons is that the latter are reasons for being in a certain affective state, that is for feeling a certain way. To believe an evaluative proposition, even the proposition that an object merits a certain feeling, is not in itself to respond to affective reasons. Similarly, acting upon an object in ways that are closely related to evaluative properties the object might bear, does not involve responding to affective reasons. Responding to affective reasons involves bearing a particular relation to certain kinds of considerations, a relation distinctive of affective, as opposed to epistemic or practical, attitudes. We will see just what sort of relation this is.

4.4.1 The Received Dichotomy

Reasons for belief and reasons for action count in favour of beliefs and actions respectively, but they do so in different ways, given that they support different types of attitude. I will take the attitudes that epistemic reasons stand in support of to be beliefs, while taking practical reasons to stand in support of actions or intentions to act[90].

Many speak of epistemic and practical reasons as distinct, and often exhaustive, types of reasons[91]. This dichotomy is supported by distinctive features that characterize each type of relation. Four distinctions are typically made between epistemic and practical reasons, they relate respectively to their: formal object, relation to value, sufficiency, and what I will call

---

[89] Raz, (1999) and Skorupski (2010) are exceptions to this trend.
[90] I follow Scanlon (1998) in using the terms 'actions' and 'intentions' interchangeably.
[91] Skorupski (2010) makes this point as well.

'binarity'. I outline the received distinctions between the epistemic and practical realms in turn, before turning to what to make of affective reasons. As I am taking reasons to be facts that count in favour of certain attitudes, reason-relations will depend on the type of attitude the reasons stand in support of. By 'types' of reasons then, I mean distinctive ways in which facts count in favour of attitudes. This allows the same fact to be an epistemic and practical reason, depending on what type of attitude the reason is counting in favour of.

Practical and epistemic reasons are thought to relate to distinct formal objects. Reasons for belief all stand in the epistemic reason-relation due to their connection to truth, which is taken to be the formal object of belief. All reasons for belief, therefore, relate to one concern, truth. Reasons for action, on the other hand, count in favour of actions in virtue of the value of performing that action. Raz (1999) writes that 'reasons are facts in virtue of which (…) actions are good in some respect and to some degree' (23). The formal object of intentions, therefore, is the value or goodness of the action in question. This can be cashed out in terms of the action's relation to 'the Good', or by citing the specific values, or forms of goodness, that discrete actions hold for the agent. Either way, epistemic reasons and practical reasons stand in support of their respective attitudes by virtue of bearing a relation to a distinct formal object.

This distinction in type of formal object means that practical reasons can relate to a plurality of values, while epistemic reasons do not. There are many ways of being good. Practical reasons that relate to different specific values can be weighed against each other in support of actions or intentions to act. In the epistemic case, all epistemic reasons count in favour (or against) the same concern: truth. When deciding whether to spend the evening reading a book, one weighs practical considerations that relate to distinct values: for example, the value of immersing oneself in literature, the value of immersing oneself in the relevant topic, one's duty to spend that evening with a friend instead, and so on. In coming to believe that the relevant book was written by a writer under a pseudonym, on the other hand, considerations of historical evidence, similarities to the writing style of other work of the same authorship, and expert opinion, all bear on the truth of one's belief, rather than on any other concern.

The difference in formal object between the epistemic and practical realms highlights the differential relation these attitudes bear to value. Practical reasons relate to consequentialist considerations while epistemic reasons do not. The value of holding a certain belief is

independent to one's reasons for having the belief. It might be beneficial for me to believe my cat to be alive, but that this is so is no reason, in the normative sense, for me to hold that belief. In the practical case, the value of taking certain actions is exactly what counts as a reason for a given action or intention. This is the case even when reasons for action relate to values we think of in non-consequentialist terms, such as duties and principles. Epistemic reasons are independent of such values as well. One might be committed to patriotic principles and duties, for example, but these are not adequate reasons to believe one's country to be innocent of war crimes. Epistemic reasons then, are independent of the consequentialist and non-consequentialist value of holding the relevant belief, while practical reasons are dependent on precisely the value of pursuing given actions.

The difference between how epistemic and practical reasons relate to value brings forth a further distinction between them, what Raz (2011) calls the 'presumptive sufficiency' of reasons. The thought is that practical reasons seem to be presumptively sufficient for intentions, while epistemic reasons are not. Let's assume we have only one reason of each sort, in each case, to illustrate the distinction. If one has only one practical reason for an action, and no further reasons for or against it, one is typically justified in pursuing that action. This is because the practical reason speaks in favour of the value of pursuing that action. Things are different in the epistemic case. In having one epistemic reason for P, while lacking any further reasons for or against it, one might not be justified in forming the belief that P. This is because the epistemic reason might not be *sufficient* for belief. One may have reason in favour of believing a picture to have been taken in Portugal, that the picture features a beach, for example, but this reason is, on its own, insufficient to justify the belief that the picture was taken in Portugal. On the other hand, if one has reason to go to Portugal, in the absence of any other practical reasons, one would be justified in going to Portugal based only on that reason.

Epistemic reasons can be better or worse reasons for thinking a certain proposition true, and therefore an epistemic reason for P might be insufficient to justify a belief that P. While one can also have better or worse practical reasons for an action, so long as one has practical reason for an action, one has a reason to value taking that action, and therefore one has reason that could, on its own, justify the action. I call this feature 'sufficiency', but I have in mind the same meaning of this term as Raz does when he uses the term 'presumptive sufficiency'. That is, the sufficiency in question is not sufficiency in the strict logical sense

where if one reason to Φ, then one's intention to Φ is justified. Rather, the notion I am employing is that if there is a reason to Φ, this is sufficient, granted there are no intervening reasons, to justify one's intention to Φ.

This difference in sufficiency relates to the fourth relevant distinction between epistemic and practical reasons, which I called 'binarity'. The received view is that practical reasons demonstrate binarity while epistemic reasons do not (Skorupski 2010; Raz 2011). Let's continue to assume we have only one reason of each sort in each case, to illustrate the distinction. The thought is that, when one has a practical reason to Φ, one either acts in accordance with it, or does not. One either forms an intention to Φ, or does not. By not Φ-ing, or not intending to Φ, one is not responding to the practical reason to Φ. To respond to the practical reason to Φ, one must form the relevant intention. As practical reasons speak to the value of actions, in failing to form the relevant intention one is not recognizing the value of the relevant action. Practical reasons therefore display binarity in the sense that one either responds to them by forming the relevant intention, or one does not.

In the epistemic case things are different, for one can suspend belief without this meaning that one is not responding to the relevant epistemic reason. One may have reason for believing P and fail to believe P. Failing to believe P, however, does not amount to believing ¬P. Recall that we are assuming there are no other epistemic reasons at play, so the subject has only one reason, in favour of P. Responding to this reason can involve forming the belief that P, or it could involve suspending one's belief, until one encounters further evidence for P, for example. Only in believing ¬P is one not responding to one's epistemic reason for P. Epistemic reasons, therefore, do not display binarity, for there are three options, or occupiable stances, one can take towards epistemic reasons, as opposed to only two in the practical case.

The difference in binarity between epistemic and practical reasons mirrors their difference in sufficiency. As epistemic reasons are not always sufficient for the justification of beliefs, one can respond to a reason for P with suspension of judgement. As practical reasons are sufficient to justify actions, however, one either forms the relevant intention to act, or one fails to respond the practical reason.

That practical reasons display 'binarity', while epistemic reasons do not, is often taken to underlie an additional difference between the epistemic and practical realm: that of rational failure[92] (Skorupski 2010; Raz 2011). The thought is that in failing to Φ, or in failing to intend to Φ, one is irrational, given that one is not responding to one's practical reason. In the epistemic case, only if one believes $\neg P$, when one has reason to believe P, is one irrational. The thought is that suspension of belief is not irrational, as one is still responding to one's epistemic reasons. Failing to Φ, or to intend to Φ, on the other hand, is a rational failure, because in doing so one is not responding to one's practical reasons.

However, once we consider cases in which there are sufficient reasons of both type, the differences in rational failure seem to dissolve. For one, it seems that suspensions of belief, when one has sufficient reason to believe P, can involve failures of rationality (suspending belief in man-made climate change for example). But it seems that believing, or acting, against sufficient reason won't guarantee a rational failure. One might have sufficient reason to Φ, in cases where not Φ-ing would not be a rational failure (one can have sufficient reason for one action, while having an overriding reason for another). The same is true of the epistemic realm. One can have sufficient reason to believe that P, yet have overriding reason to believe $\neg P$. Rational failure therefore seems to involve considerations beyond sufficiency, namely, attention to whether, and what, other reasons are at play. It, therefore, seems strange to count differences in rational failure as a further distinction between practical and epistemic reasons, given that in cases of sufficient reason of each sort, we observe similar patterns of rational failure in both practical and epistemic realms. It has been my interest to highlight the received review on the differences between practical and epistemic reasons, without committing myself to particular ways of understanding details within them, or how they feed in to other concerns. I will, therefore, remain tentatively agnostic regarding differences in rational failure between the two realms, and move forward with the four distinctions characterized above.

4.4.2 The Irreducibility of Affective Reasons

Responding to an affective reason involves feeling a certain way in virtue of certain properties of the object. In section 4.3 I outlined empirical work in support of this characterization. We

---

[92] Here 'rational failure' means irrationality in the reason-responsiveness sense. That is, it pertains to failures in responding to reasons, rather than to failures in being guided by norms of rationality.

saw that emotions involve taking features of their objects to count in favour types of appraisal. I will argue that there is a distinctive type of reason-relation between emotional attitudes and facts that count in their favour. This is because affective reasons do not stand in support of emotions in ways that are reducible to practical or epistemic reason-relations. Let's look at how affective reasons do regarding the differences between the received dichotomy of practical and epistemic reasons, namely regarding; formal object, relation to value, sufficiency, and binarity.

If emotions, as a kind, admit of one formal object, it is not likely to be truth. Most contemporary philosophers use notions of appropriateness, fittingness or aptness, as opposed to truth, in speaking of the correctness conditions emotions[93]. One reason for doing so, is that emotions are not thought to involve propositional attitudes that some deem necessary for truth-aptness (Salmela 2014: 105). Another reason is that, by often involving conative components, with world-to-mind direction of fit, emotions seem to be characterized by a type of intentionality that is not reducible to the mind-to-word direction of fit of beliefs, which is most amenable to truth-aptness. A clear reason to move away from truth as the relevant aim of emotional attitudes, is that truth often comes hand-in-hand, with objectivity. There can be facts that stand in support of your sadness, while they do not stand in support of mine, your dog having passed away for example. Reasons for emotion do not necessarily speak indiscriminately in favour of all agents having the relevant emotion. On the other hand, that there be reason to believe water is composed of H_2O molecules, is an epistemic reason in favour of this belief for everyone. Appropriateness can be seen as a truth-analogue in the emotional realm that captures these distinctions[94]. Appropriateness is therefore an intuitive candidate formal object for emotions, as opposed to truth.

Affective reasons count in favour of the appropriateness of particular emotional attitudes. Emotions, as a kind, are not often construed as being related to one formal object, however. It is particular emotion types that are thought to have formal objects; danger for fear and offence for anger, for example (Deonna and Teroni, 2012). If this is the case, then affective

---

[93] Salmela (2014) and de Sousa (2007) are exceptions to this trend. See Salmela (2014: Ch.5) for a defense of truth-aptness being the relevant assessment for emotions. Note that even if their arguments go through, and emotions are shown to share the same formal object as beliefs, this does not, on its own, challenge my argument that affective reasons are irreducible to epistemic ones. This is because, as we shall see, affective reasons will still differ from epistemic ones on account of both sufficiency and binarity.

[94] If an account of emotions as truth-apt can capture such distinctions, then I am happy to grant that emotions aim for truth. This would, however, make the issue mostly a terminological one, as emotional truth-aptness would still bear differences to truth-aptness in the epistemic realm.

reasons differ from epistemic reasons by virtue of not relating to one particular formal object. On either construal of the formal object of emotions then, affective reasons and epistemic reasons relate to distinct concerns.

In contrast to the practical realm, where practical reasons can be thought of as counting in favour of the Good of pursuing a particular course of action, affective reasons do not count in favour of the Good of feeling a certain emotion. Many emotions are not good to feel, they do not feel pleasant nor do they have beneficial causal effects for the agent. Perhaps affective reasons are more akin to practical reasons construed as relating to distinct values. After all, we saw above that practical reasons need not be thought to relate to the Good, but can be thought of as relating to the specific values of distinct actions (Raz 2011).

Reasons for emotions might be thought to relate to specific evaluative properties much like reasons for action relate to the different values of specific actions. A key difference between the affective and the practical realm is that specific values do not separate actions into action types however, while evaluative properties are thought to individuate emotion types. This means that emotion types will typically be supported by affective reasons that relate to one type of evaluative property or formal object: danger for fear, offence for anger, for example. A single action, on the other hand, can derive support from an open-ended range of values: that the action will be aesthetically pleasing, that it will bring about personal satisfaction, that it will solidify your friendship, for example. With affective reasons this is not the case. A single emotion derives support from affective reasons that count in favour of a particular evaluation. Only considerations relating to the bear's dangerousness count in favour of fear, for example.

This isn't to say that affective reasons that count in favour of distinct evaluations cannot support mixed emotional states. Nostalgia and thrill, for example, seem to involve a mixture of evaluative properties. There is, however, a limit to the range of distinct sources of evaluative support that can count in favour of mixed emotional states, whereas the range of distinct values that can count in favour of one particular action is entirely open ended.

We, therefore, have reason to think that emotions do not share the formal object of action or belief. This is unsurprising on any account that takes emotions themselves to be irreducible to beliefs or actions. Affective reasons are perhaps best construed as counting in

favour of the appropriateness of the formal object of distinct emotion types. That is, affective reasons count in favour of certain evaluations being appropriate. The aim that all emotion types share is that of being an appropriate evaluation, while the type of evaluation differs depending on emotion type. This fits well with the distinctions outlined in this section, as well as with the role I have argued attention plays in highlighting and 'checking' features of objects to count in favour of particular emotional evaluations. Affective reasons are, therefore, typically properties of the particular objects of emotions, that count in favour of the attribution of certain evaluative properties[95].

What about the relationship between affective reasons and the value of the attitudes they stand in support of? We saw that epistemic reasons are not reasons that concern the value of holding certain beliefs, but merely concern the truth of the belief. Epistemic reasons are ones that are *followed* in forming the relevant belief. Practical reasons for holding beliefs, such as it being beneficial to an agent that they hold a certain belief, are not reasons one can follow in the formation of beliefs. They do not justify beliefs, although they might motivate beliefs by justifying actions an agent can take to feign belief in them. We saw that practical reasons, on the other hand, are precisely concerned with the value of performing, or taking, certain actions. On this point, affective and epistemic reasons are much the same. Proper normative reasons for emotions are justifying reasons for the emotional state itself, as opposed to reasons that might count in favour of bringing a certain emotional state about. Affective reasons, therefore, do not concern the value of experiencing certain emotions, as this is orthogonal to whether certain affective reasons hold. That becoming angry at your friend would cause them to give you their pocket money, for example, is not a reason that can justify your anger in the normative sense. Such a consideration could justify your taking actions to become angry, or to feign anger, but it could not justify your anger itself.

The similarity between affective and epistemic reasons does not extend to the third distinction: sufficiency. With respect to sufficiency, affective reasons seem more akin to practical reasons. Recall that one epistemic reason is not necessarily sufficient to justify a relevant belief, while one practical reason is. An epistemic reason can count in favour of the truth of a belief, without justifying it. On the other hand, if one has a practical reason, one

---

[95] I said in fn.47 that I favoured a form of value realism where evaluative properties supervene on non-evaluative properties, whether these evaluative properties are response-dependent, or in-dependent. Affective reasons then, can be seen as the non-evaluative properties on which response-dependent, or response-independent, evaluative properties supervene.

has enough reason to justify taking the relevant action. Affective reasons count in favour of attributions of particular evaluative properties to the objects of emotions. If one has an affective reason for a certain emotion then, one has reason to attribute the object of one's emotion the relevant evaluative property. One affective reason is therefore sufficient to justify the relevant emotion. That we often form beliefs that conflict with our emotions speaks in favour of this picture. One might not have sufficient reason to believe a new acquaintance is untrustworthy, but one might have reason to feel suspicious, if there are features of one's acquaintance that count in favour of suspicion, given that affective reasons are presumptively sufficient.

What about binarity? Recall that one can occupy three distinct stances on one's epistemic reason for P; belief in P, suspension of belief, and belief in ¬P. Only believing ¬P involves not responding to one's epistemic reason for P. As opposed to the epistemic case, the practical case is characterized by binarity. One either acts (or intends to act), in response to one's reason, or one does not. We might think affective reasons are not characterized by binarity, after all, if there is reason for anger there are more than two possible outcomes, one can: feel anger, not feel anger, or, one can feel a different emotion, let's say joy, for example. This suggests that affective reasons are analogous to epistemic reasons, as opposed to practical reasons, when it comes to binarity[96].

I think the analogy with the epistemic case here might be too quick however. Suspending belief is not clearly analogous to 'failing to feel'. In suspending belief, one remains uncommitted to a belief in P, and one does not endorse ¬P. By failing to feel anger, when one has reason to feel so, and no further considerations at play, one is not 'suspending affect' (if sense can be made of such a notion), one is simply not feeling anger. The first disanalogy, therefore, arises from the difficulty of making sense of what 'suspension' of feeling would involve. A second disanalogy arises from the fact that there is also no clear analogy to believing ¬P in the emotional realm. Is anger's opposite really joy? One can imagine people deriving Machiavellian delight in having been slighted, as they now get to plot revenge.

Even in common instances of anger, anger has not been characterized as an entirely negative emotion, as we saw that it includes some optimism for change within it. Pairing emotions with opposing valences together will, therefore, not only be difficult, but, crucially, doing so

---

[96] Raz (1999) and Skorupski (2010) seem to think this to be the case.

will not ensure that these pairs are 'opposites'. Furthermore, seemingly 'opposite' pairs of emotions do not seem to contradict each other. For example, it would not be necessarily contradictory, to feel anger at someone but also compassion for them. Perhaps they have harmed you in a way that demonstrates their own self-loathing. So long as there is reason for compassion, the fact that one feels compassion, when there is also reason for anger that one fails to respond to, does not amount to feeling something opposed to anger. The failure seems to be in not responding to reasons for anger, rather than in responding to reasons for a distinct emotion.

What can occur, is that one can experience an emotion that is inappropriate, that is, an emotion for which there are no affective reasons. If one felt joy when there was no reason for joy, for example, then one's joy is inappropriate. But this would be analogous to holding a separate belief, Q, on false or insufficient grounds, rather than being analogous to holding ¬P, when one has reason to believe P.

Provided one has affective reason for anger, one has reason to feel that an offence has occurred. By not feeling anger, one is not responding to a reason to feel that an offence has occurred. There doesn't seem to be a third space to occupy in response to affective reasons, one either responds to them, or doesn't. This seems analogous to the practical case where one either acts, or does not act, on one's reason for a particular action. Feeling a separate emotion, for which one has separate reasons, or for which one lacks reasons, does not directly bear on one's reasons for anger, and therefore doesn't constitute a separate stance one can hold towards one's reasons for anger. Take the practical case for example, one can perform a distinct act, Θ, for which one has separate reasons, or no reasons at all, without thereby violating binarity regarding one's reason to Φ. If one responded to reasons to Φ by Θ-ing, one is still not violating binarity, one is performing an action for which one has no reason. The same is plausibly true if one responds to reasons for anger with joy, in the absence of reasons for joy. Affective reasons, therefore, seem to demonstrate binarity. Indeed, this is in line with them demonstrating sufficiency, for if one affective reason is enough to justify an emotion, then only two options seem available in response to such reasons. Due to these disanalogies, I think it is, at the very least, unclear where affective reasons stand concerning binarity.

What does this mean for rational failure in the case of emotion? Typically, failing to feel is not seen as a rational failure (Raz 2011). That a Buddhist monk faces offence unemotionally, is not taken to be a rational failure on their part, and rational demands to feel are not typically endorsed. Despite this, it is acknowledged that there will be some cases where failures to feel constitute rational failures. Raz (2011) for example writes that:

> For the most part even when appropriate there is nothing amiss in not having it. Fear is appropriate when facing great danger, but only exceptionally would its absence when in danger be against reason. (49)

The picture I have painted of affective reasons so far, suggests that such cases might not be so rare. After all, if not responding to sufficient reason for belief or action, in cases where there is no overriding undermining reason against these reasons, often involves rational failure, why should the same not apply to affective reasons in similar conditions?

One reason not to see failures to respond to affective reasons as rational failures, is that such cases are often seen as moral rather than rational failures (Raz 2011; Worsnip 2016). Failing to feel anger at social injustice would presumably count as a moral rather than rational failure. Although I will not pursue such an argument here, there seems to be room to argue against all failures to feel constituting moral, as opposed to, rational failures. First, it is unclear why we would should think of moral and rational failures as mutually exclusive. Second, there might well be cases of rational failure that are not cases of moral failure. Failing to feel suspicious of a friend that has betrayed your trust over and over again, for example, seems like a rational failure.

An additional reason for why failures to feel might not typically be seen as rational failures, is that emotions are not the only way of attributing evaluative properties to objects. One can hold an evaluative belief that one's friend is untrustworthy without feeling suspicious. Perhaps, then, we are reluctant to take failures of feeling to be rational failures, so long as one is responding non-emotionally to the relevant evaluative evidence. This doesn't, however, preclude that failures to feel might constitute rational failures. Firstly, cases where one fails to form the relevant evaluative belief *and* fails to feel the relevant emotion, seem to involve greater rational failure than cases where one experiences the emotion but lacks the belief. This suggests that failing to feel is, in some sense, a rational failure. Secondly, we might think that in a case where an agent has an evaluative belief, but lacks an associated emotion, that there is something disingenuous, or at least distinct, about the way they evaluate the

situation at hand. If someone believes that the recent killings of peaceful Somalian protestors was wrong, but fails no feel anger, we might think there is some condemnable failure here, despite their accurate epistemic evaluation. More needs to be said about what sort of deficiency is involved in failing to feel, where this deficiency is not restricted to moral failure. For now, I hope to have highlighted that much like with any reason-responsive attitude, there is room to argue for cases of rational failure when one fails to respond to affective reasons.

I have argued that reasons for emotions are irreducible to practical or epistemic reasons. We saw that affective reasons have distinct formal objects to both practical and epistemic reasons. We saw that affective reasons are like epistemic reasons only with respect to being independent from the value of holding the relevant attitudes. With respect to sufficiency, we saw that affective reasons are best construed in line with practical reasons, that is, as being presumptively sufficient to justify emotions. With regard to binarity things are less clear, but I have provided reason to think affective reasons are also in line with practical reasons in this regard. Affective reasons are characterized by a set of features that is distinct from the set of features that characterizes practical or epistemic reasons. This suggests that affective reasons are irreducible to epistemic or practical reasons.

4.4.3 Strictness

I have carved out a reason-relation that holds between facts and the emotions they count in favour of. We saw that this reason-relation is best construed as sui generis in that it is irreducible to epistemic or practical reason-relations. There is a feature distinctive of the emotional realm, that deserves its own label. This is a feature that has been at play in the discussion above, but not characterized, that bears on interactions between reasons. The feature is the following: affective reasons do not cancel each other out. This is because reason for one emotion does not count against reason for another. In not cancelling each other out, affective reasons are, what I call, 'strict'.

We saw this in the case where there was reason for both anger and joy. That there be reason for two distinct emotions, which might seem, and even feel, to be in tension, does not mean that the affective reasons cancel each other out. On the contrary, affective reasons for distinct emotions provide independent support for their respective emotion. This fits with what I have said about emotions not having clear opposites, as well as with each emotion having its

own formal object. In the case of belief, epistemic reasons can cancel each other out in support of a belief, as their only concern is truth. Similarly, in the practical case, practical reasons, relating to an open-ended range of values, can outweigh and cancel each other out in support of the value of a relevant action. With emotions, affective reasons stand in support of a specific emotion, and, therefore, do not directly compete with affective reasons for distinct emotions.

Strictness is a distinctive feature of affective reasons, as both epistemic and practical reasons can outweigh each other, or cancel each other out. We have seen that practical reasons that favour opposing actions can cancel each other out. One's reason to go to a party tonight (that it is one's friend's birthday) can be cancelled out, or outweighed, by a reason not to go (that one is not feeling well). Similarly, in the epistemic realm, conflicting epistemic reasons can cancel each other out. Reason to believe your friend is pregnant (that she's been feeling nauseous) might be outweighed by reason to think she is not (that she got drunk last night). Things are different with affective reasons. That there are affective reasons for one emotion does not cancel out affective reasons there might be for other emotions. Affective reasons are in this sense strict[97], they provide support for emotional attitudes even when there are other affective reasons at play. Affective reasons are, therefore, strict in part because they count in favour of the appropriateness of the formal objects of particular emotions. By counting in favour of the appropriateness of distinct formal objects, reasons for fear do not cancel out reasons for joy, for example.

Affective reasons can, however, much like epistemic and practical reasons, combine in support of the same attitude, there may be multiple reason for sadness, for example. One's

---

[97] Maguire's (2018) uses the term 'strict' to refer to facts that are 'individually normatively potent', and that are not weighed against others in support of relevant attitudes. I use the word in a similar but distinct sense. For me, 'strict' refers to an inability to be cancelled out by competing reasons of the same sort. Therefore, my notion of strictness, as opposed to Maguire's, allows affective reasons to be play some weighting roles but not all. I think affective reasons can combine in support of emotions, that some affective reasons can be stronger than others, and that one affective reason can affect the force of another. I do not, however, think that affective reasons can outweigh or trump each other, that is, they cannot completely cancel each other out. This is the sense of strictness I employ. Maguire denies that facts that stand in support of emotions play any weighting roles at all. He does not think they can combine in support of emotions, nor that some provide stronger support for emotions than others, nor that they can affect each other. Maguire argues that these features of emotion-supporting facts makes them fall short of reason-hood. He thinks emotions are subject to fittingness conditions, rather than reason-relations. His argument against emotions being had for reasons faces a number of problems, not least the underdetermination of the alternative standard of normative support that emotions are meant to be governed by. In addition to this, Maguire's working conception of reason-hood seems to exclude beliefs from having reasons as well, which suggests he is starting from too narrow a conception of the reason-relation. For a reply to Maguire see Faraci (2018).

reappraisal mechanisms can adjust to the presence of previously unnoticed reasons for a particular emotion. Affective reasons can also combine in support of mixed emotions such as thrill or nostalgia. Affective reasons cannot, however, cancel each other out. When there are reasons that count in favour of distinct emotional appraisals, either a justified mixed emotional state will result or one will have reason to feel either one of the emotions. Indeed, the view that it is often warranted to feel conflicting emotions is widespread in the philosophy of emotion literature (Greenspan 1988; Tappolet 2005). Strictness is an important feature of the normativity of emotions, as it underlies distinctive features of our emotional life, such as ambivalent or mixed emotions.

Mixed emotions are particularly prevalent for those living under conditions of oppression, where subordinated agents tend to develop a 'double-consciousness' of which distinct and often conflicting emotional experience arise. Under oppression, agents have a split view of reality, and themselves. One view reflects internalized dogma and the other allows a critical stance on the first, as well as the resources to imagine a different reality. Du Bois (1996) writes of the African American experience: 'One ever feels his twoness, - an American, a Negro; two souls, two thoughts, two unreconciled strivings; two warring ideals in one dark body' (215). The two consciousnesses naturally give rise to distinct emotional responses to the same state of affair. Both might be justified by the affective reasons available. The strictness of affective reasons means that these conflicting considerations, that stand in support of distinct emotional attitudes, do not cancel each other out.

Strictness also underlies an important difference between responding emotionally to affective reasons and responding epistemically to reasons that favour evaluative beliefs. The same fact can count in favour of a belief, and an emotion, but the fact is governed by a distinct reason-relation in each case. When there is reason to find a dog dangerous, this fact acts as an affective reason, in support of the emotion of fear, and/or as an epistemic reason, in support of a belief featuring an evaluative concept regarding the dog's dangerousness. Responding to a reason to find the dog dangerous affectively, will involve feeling fear, while responding to the same reason epistemically, will involve forming an evaluative belief about the dog's dangerousness. We can call evaluatively significant reasons 'evaluative reasons' for ease. Evaluative reasons don't carve out a distinct reason-relation, though; they are facts that count in favour of attitudes with evaluative content. Evaluative reasons are governed by a different reason-relation depending on which attitude they count in favour of. All affective reasons

are evaluative reasons, but evaluative reasons can also act as epistemic reasons for evaluative beliefs, and as practical reasons for actions as well[98].

Because affective reasons are strict, however, other affective reasons will not cancel out reasons for fear, they will just support distinct emotions. As epistemic reasons are not strict, if there are evaluative reasons in favour of forming distinct evaluative beliefs, these can cancel out the evaluative reasons for forming the belief that the dog is dangerous. If there are reasons to find the dog exciting for example, these can cancel out evaluative epistemic reasons to believe the dog is dangerous. In the emotional case this is not so, as affective reason to feel excitement cannot cancel out an affective reason for fear. This fits well with the fact that we often feel emotions without believing the associated evaluative belief. On my account, the strictness of affective reasons allows them to justify emotions despite there being other affective reasons at play. The very same facts might not be able to justify an evaluative belief, however, if there are separate evaluative considerations that outweigh it.

Strictness is distinct from sufficiency. Sufficiency says that one affective reason is sufficient to justify the relevant emotion. Sufficiency does not say that affective reasons cannot be undermined by further affective reasons. 'Strictness' is meant to capture this, as it tells us that affective reasons don't count against each other. Strictness captures the fact that one affective reason cannot be defeated by another, making sufficiency true of cases where there are multiple affective reasons for distinct emotions. This means that if there are two affective reasons, each in favour of a distinct emotion, they would both be sufficient to justify their respective emotion, given strictness and sufficiency.

We might think sufficiency extends to even more cases than those accounted for by strictness. Perhaps affective reasons are sufficient to justify relevant emotions, even in the presence of not only further evaluative reasons, but also some non-evaluative epistemic reasons too. That there be some epistemic reason that counts against an affective reason one has, does not seem to necessarily undermine the affective reason. This is because, as affective

---

[98] Indeed, all practical reasons might be evaluative, in that they concern the value of pursuing certain actions. But specifically, the types of evaluative properties that feature as the formal objects of emotions can count in favour of particular actions because it is good to respond to them practically in a particular way. Many facts can act as reasons for distinct types of attitude. This is not solely a feature of what I have called 'evaluative reasons' for brevity. Facts pertaining to food can act as reasons to seek out food, as well as reason to believe some food is one's favorite for example. We can carve out many classes of reasons that relate to distinct features of the world, this does not mean that there are distinct reason-relations characteristic of each class, however. Reason-relations are relations between particular types of *attitude*, and facts that count in their favour.

reasons are individually sufficient for the justification of emotions, we might think only *sufficient* epistemic reason is able to undermine an affective reason. That there be an epistemic reason to believe a vicious dog is not dangerous (that it is tied up, for example) may not be sufficient to undermine one's affective reason for fear. If one has *sufficient* reason to believe the dog is not dangerous (that it is tied up by a chain made of an indestructible metal, for example), these epistemic reasons might be enough to undermine the justification of one's fear. The same is not so in the epistemic realm. Epistemic reasons do not demonstrate sufficiency, therefore, having one evaluative reason for a belief is not always sufficient to justify the belief. This means that, an evaluative reason in favour of the dog's viciousness might not be sufficient to justify the belief that the dog is dangerous. This, in turn, means that one's evaluative reason to believe the dog dangerous can be undermined by separate epistemic reason (that the dog is tied up). That is, one might not need *sufficient* epistemic reason (that it is tied up by a chain made of an indestructible metal) to undermine an evaluative reason in favour of a belief, but merely some epistemic reason to do so. As epistemic reasons do not demonstrate sufficiency, one epistemic reason that conflicts with one's evaluative epistemic reason for a belief can be enough to undermine one's evaluative epistemic reason.

We, therefore, have reason to think that emotional justification is more resistant to wider considerations than epistemic justification. This is because affective reasons can count in favour of certain emotions, even in the face of any other affective reasons, as well as in the face of some epistemic reasons. Epistemic reasons do not. In the next chapter we will see how strictness and sufficiency underlie the distinctive epistemic role emotions can play as reasons for evaluative beliefs. For now, I hope to have characterized some plausible and important features of the sui generis reason-relation distinctive of emotions.

4.5   Conclusion

I have argued that emotional reason-responsiveness is best construed as *basic*, that is, as independent to conformance with norms of rationality, and as *sui generis*, that is, irreducible to epistemic or practical reason-responsiveness[99]. I began by distinguishing two ways in

---

[99] I did not address the fact that emotions, as opposed to other reason-responsive phenomena, are assessible along the dimension of proportionality. That is, two situations that call for anger may call for distinct intensities of the emotion. The account of affective reasons that I proposed, however, might have the resources to account

which emotions have been considered irrational, one pertaining to the violation of norms of rationality, the other concerning emotions as inappropriate, or unresponsive to reasons. I described two types of account that can be provided in order to make sense of emotional reason-responsiveness, one that takes reason-responsiveness to depend upon compliance with requirements of rationality, and the other which takes the reason-relation between facts and attitudes to be basic. I argued that the latter approach should be preferred despite it being the least developed in the literature. I outlined Agential Disposition accounts, which represent the most fleshed out accounts of emotional reason-responsiveness in the literature, and argued that they do not succeed in straightforwardly avoiding key problem cases.

I then characterized emotions as phenomena capable of responding to reasons, rather than merely tracking them. The attentional dynamics at play during emotional episodes support my construal of emotions as involving evaluative stances towards objects, where features of these objects are experienced as counting in favour of the emotion. I then proceeded to characterizing this basic reason-relation of the emotional realm, and argued that there is reason to endorse a trichotomy of reason types[100], in place of the traditional dichotomy of reasons that are either practical or epistemic. We have reason to think reasons for emotions carve out an irreducible type of reason-relation, where strictness and sufficiency emerge as key features. Anyone arguing that reasons for emotions, or affective reasons, are reducible to epistemic or practical reasons must explain how the traditional dichotomy of reasons can account for features of our emotional life I have highlighted, or explain why we should view some of these features as not properly supported by reasons.

I have, therefore, taken significant steps towards fulfilling my third desideratum regarding the provision of an account of anger as reason-responsive, or normatively rational. I have

---

for proportionality assessments. Firstly, different affective reasons can have different weights, much like epistemic and practical reasons can have different weights. That your friend harmed you is reason for anger, but that they harmed you out of carelessness might be reason for less intense anger than if they had harmed you with perverse intent. Affective reasons can also combine in support of more intense emotions. That you received a desirable job offer, and were awarded your doctorate, on the same day, is reason for more intense happiness than if you had received only one of the two. Note that affective reasons having different weights, and being able to combine, is compatible with them being strict. Recall that in being strict, affective reasons for distinct emotions cannot cancel each other out. This does not mean that affective reasons cannot stand in weightier than relations to one another. That reason for joy not cancel out reason for anger, does not mean that one cannot have *more* reason for joy than anger. To say that an emotion is reason-responsive, then, might include considerations of proportionality, as to respond to one's affective reasons may involve experiencing an emotion at the intensity for which there is reason.

[100] See Skorupski (2010) for an argument for the irreducibility of evaluative reasons that bears similarities to my own.

argued that a particular approach to emotional reason-responsiveness is to be preferred and I have taken important steps in building an account of this type.

On my account, emotions can be reason-responsive despite violating rational requirements for coherence or enkrasia. Emotional reason-responsiveness is basic, in that it is independent to concerns regarding the downstream effects of emotions. This is important because it carves out an account of the normativity of emotions that distinguishes between distinct types of failure that would otherwise be obscured: failing to have appropriate emotions, and failing to comply with what have been called 'rational requirements'. A basic account of emotional reason-responsiveness also allows us to attend to the intricacies of the reason-relation characteristic of emotions. These would be harder to make out if reason-responsiveness were cashed out in terms of compliance with general requirements of adequate reasoning. Whether emotions actually tend to undermine rational requirements is an empirical question, the answer to which should not be assumed a priori[101], and which must be properly weighed against the positive epistemic role emotions play in responding to reasons. I suspend further discussion on this until after I have tackled anger's epistemic rationality. For now, I note that in cases of rational akrasia, and outlaw emotion, the reason-responsiveness of emotions is crucial. A basic account of reason-responsiveness allows us to make sense of these cases as ones where the emotions are reason-responsive despite requirements of rationality being violated. This suggests that emotional reason-responsiveness is, at least sometimes, more valuable than compliance with rational requirements.

---

[101] Blanchette et al. (2014) provide a helpful review of the evidence typically cited in support of the deleterious effects of emotion on logical reasoning, and present their own contrasting results.

5

Emotions as Reasons for Belief

'Recognizing certain neglected aspects of emotion
makes possible a better and less ideologically biased account
of how knowledge is, and so ought to be, constructed.'
-Jaggar, in 'On Love and Knowledge'

In this chapter I am concerned with the epistemic role of emotions. Our fourth, and final, desideratum is to provide an account of anger's epistemic rationality. Anger's epistemic rationality concerns whether it has a positive effect on leading agents to form beliefs based on adequate evidence. Those who condemn anger take it to be epistemically irrational because, even if anger can pick up on reasons for it, they take anger to be an unreliable, and more epistemically costly, way of doing so, as compared to beliefs. This is because they take anger to generate more false-positives than beliefs, and to bias the formation of beliefs in favour of the agent's interests[102]. Although those philosophers that have praised anger take it to be epistemically rational, as they think anger plays not only

---

[102] We saw this in chapter 1, sec 1.1.

positive, but crucial, roles in leading agents to gain knowledge, their arguments in favour of this remain largely programmatic. This chapter will take steps to remedy this.

To meet my fourth desideratum, we must consider the epistemic role played by emotions as a mental kind. Work on the epistemology of emotions can help here. I will argue that emotions can play direct justificatory roles in support of evaluative beliefs. This is a thesis popular in the philosophy of emotion. It is, however, one in need of defence, given my commitment to emotions admitting of justification themselves, as we will see. That emotions play justificatory roles themselves might not be sufficient for anger to enjoy epistemic rationality, after all, anger might have noxious epistemic effects to an extent that they undermine its positive ones. If emotions are epistemically necessary for the justification of some beliefs however, this sets the bar much higher for those seeking to condemn anger on epistemic grounds. That is, if anger is necessary for the justification of certain beliefs, it can play an indispensable epistemic role that eclipses many potential noxious epistemic effects. I will argue that anger, indeed, has such an indispensable epistemic role. This will be made clear in my defence of emotions playing justificatory roles.

In the philosophy of emotion literature, the debate regarding the epistemic role of emotions centres around popular perceptual theories[103] and whether the strong epistemic role they afford emotions is plausible. Perceptual theories take emotions to be analogous to perceptions in a number of ways[104]; both allow us to apprehend objects in the world by virtue of representing their objects non-conceptually, both are thought to have distinctive phenomenologies and both provide an immediate source of prima facie justification for relevant beliefs[105] (de Sousa 1987; Tappolet 2016; Döring 2007).

The epistemic analogy involves taking emotions to provide defeasible justification for evaluative beliefs based on them (de Sousa 1987; Johnston 2001; Döring 2007; Pelser 2014;

---

[103] The term 'perceptual theory' has been used to refer to Jamesian theories where emotions are perceptions of bodily changes, as well as to theories that take emotions to be perceptions of value. By 'perceptual theory' I mean the latter rather than the former throughout. Prinz (2006) attempts to combine the two types of perceptual theory, see Goldie (2006) for review and critique.

[104] Perceptual theories of emotion typically assume a classic representationalist account of perception (Döring and Lutz 2015)

[105] I follow the literature in taking justified belief to be the relevant epistemic goal of emotions, at least in the first instance. If knowledge is assumed to be undefeated justified true belief, then emotions lead to knowledge when the beliefs that they justify are true and no defeaters are at play. Whether facts the agent isn't aware of count as defeaters is debated, but mental-state defeaters will defeat the justification of a prima facie justified true belief, and in doing so undermine knowledge (Moretti and Piazza 2018).

Tappolet 2016). The thought is that, much like one's perception that there is a dog in front of me for example, provides defeasible justification for my belief that there is a dog in front of me, anger at my friend provides defeasible justification for the belief that my friend has wronged me. This epistemic thesis has been called 'epistemic perceptualism' (Cowan 2016; Mitchell 2017) given its endorsement by perceptual theorists of emotion. It should be noted that this name is somewhat unsatisfactory however, as one can endorse this epistemic thesis without being a perceptual theorist and without being committed to any particular view of perceptual justification[106]. Indeed, the epistemic thesis has gone simply by the 'justification thesis' (Pelser 2014) or the 'justification view' (Mitchell 2017), as the epistemic role proposed of emotions is a normative one. I follow this trend and call it the 'Justification Thesis' (JT), given that my concern is with evaluating the positive epistemic thesis regarding emotions alone. I will not be concerned with whether the view holds true in the perceptual case, nor with whether emotions should be construed as analogous to perceptions in the various other ways perceptual theorists have highlighted[107]. I am concerned with whether, and how, emotions play a justificatory epistemic role.

I begin by outlining the Justification Thesis (JT), and why I am motivated to defend it, in section 5.1. In section 5.2 I defend JT against three objections raised against it in the literature. These are, the Dependency Objection, the Superfluity Objection, and the Objection from Unreliability. In section 5.3 I put forward an overlooked and important objection to JT: the Objection from Outlaw Emotions. Outlaw emotions, as we have seen, are ones that arise under conditions of oppression, and go against the prevailing ideology. We saw that anger was hailed as perhaps the most important of outlaw emotions as, in picking up on injustices, it is typically the first epistemic achievement in resisting oppression. It is therefore crucial that our epistemology of emotions can make sense of outlaw emotion cases. I argue that JT, as it stands, risks ruling such cases out, but I will argue, that it need not. I argue that JT must be informed by an externalist social epistemology in order to account for outlaw emotion cases. I sketch the form such an account might take in section

---

[106] Deonna and Teroni, (2012: ch.10) for example, can be read as endorsing this epistemic view, despite not identifying as perceptual theorists.

[107] Recall that it was not an aim of this thesis to defend one theory of emotion over another. In chapter 2 I outlined my commitments regarding the nature of anger: that it is an intentional state with evaluative phenomenal content, that sometimes represents its objects conceptually and other times do so non-conceptually. Given that perceptual theories can come in quite liberal forms, I don't take the question of whether my commitments align me with perceptual theories or not to be a particularly philosophically interesting one. For a detailed introduction to perceptual theories of emotion see Döring and Lutz (2015). For a critique of perceptual theories see Salmela (2011) and Brady (2011).

5.7. My aim in this chapter then, is to defend the view that emotions can justify beliefs, as well as put forward a preferred framework for making sense of this justificatory role. In doing so, a case will be made for anger playing crucial epistemic roles. This chapter will, therefore, provide an account of the epistemic role of emotions that supports anger being epistemically rational.

5.1   The Justification Thesis

Emotional experiences often give rise to beliefs. That this is the case 'reveals a human proclivity to trust emotions, at least implicitly, as justifying evidence or reasons for belief' (Pelser 2014: 108). The main epistemic upshot of perceptual theories is that we are prima facie justified in trusting our emotions. Specifically, perceptual theorists hold that emotion-based beliefs are justified much the same way perceptual beliefs are. The thought is that both emotions and perceptions allow us to become aware of certain objects and as such both provide defeasible justification for beliefs about these objects. Whereas perceptions provide prima facie justification for empirical beliefs, emotions are thought to provide prima facie justification for evaluative beliefs. Emotions, construed as perceptions of value[108], play an analogous epistemic role to the one perceptual states have been taken to play.

Whether emotions are direct experiences of values, or are based on 'cognitive bases', such as perceptions, beliefs, or memories, from whom they inherit part of their content, the thought is that emotions have evaluative content that can justify corresponding evaluative beliefs[109]. Tappolet, a contemporary perceptual theorist, writes:

> If we accept the claim that emotions have contents of this sort, then it becomes natural to claim that emotions are like sensory perceptions in that they allow us to be aware of certain features of the world, namely values. They do so, at least, under

---

[108] I said in chapter 2 (fn.47) that I follow most contemporary philosophers of emotion in assuming some form of value realism. I said I leaned towards an account where evaluative properties supervene on non-evaluative ones, where evaluative properties can then be cashed out in either in a manner that is response-dependent or response-independent.

[109] We saw that most philosophers of emotion take emotions to have evaluative content of some sort (be it conceptual evaluative content for judgement theorists or non-conceptual evaluative content for perceptual theorists). Deonna and Teroni (2012)'s Attitudinal theory of emotion is an exception, as they think that emotions have non-evaluative representational content and that the evaluative component of emotions is supplied by emotions involving evaluative attitudes. They would however presumably not deny that emotions have evaluative phenomenal content. By evaluative content I mean that emotions feel as though their objects instantiate evaluative properties. They therefore have evaluative phenomenal content. It is this phenomenology of emotions that provides evidence for evaluative beliefs according to JT. I remain agnostic on whether the phenomenological content of emotions is reducible to its representational content.

favourable circumstances, that is, when nothing interferes with them. (Tappolet 1997: 8)

Johnston (2001) writes that 'affect is akin to perceptual experience' (189), and that it can make true 'immediate (i.e. non-inferential) judgement' (205). Döring (2003) agrees that an emotion can provide immediate justification to an evaluative belief based on it, 'like a perception, it can do so by its representational content ... justifying the content of the belief' (215).

The thought, therefore, is that perceptual beliefs stand to perceptual experiences as evaluative beliefs stand to emotions. Similarly, the objects of emotions stand to emotions as the objects of perceptions stand to perceptual experiences. Perceptions of p typically give one reason to believe that P. Given this, it seems that emotions, which involve seeing a certain object x as possessing an evaluative property F, gives one reason to believe that the object x is F.

This is the epistemic upshot of perceptual theories. This upshot can be phrased in terms of emotions constituting reasons for beliefs. Emotions, much like perceptions, are typically taken to provide defeasible reason for relevant beliefs. Much like perceptions then, emotions provide defeasible reason for associated beliefs such that beliefs based on emotions are prima facie justified.

This epistemic thesis can be summarized as follows:

> **Justification Thesis** (JT): Emotions provide immediate prima facie justification for evaluative beliefs based on them.

Immediate justification is justification that is not mediated by an intermediary mental state, and that is not linked by any inferential steps. For example, your suspicion of a friend provides prima facie justification to the belief that she might be hiding something. A distinct belief regarding the trustworthiness of our emotion under the current circumstances, or any other intermediary belief, is not needed. Requiring such intermediary beliefs would mean that emotions are not capable of directly justifying evaluative beliefs. Nor is your belief that your friend is hiding something inferred from your suspicion, your suspicion is direct evidence for the belief. JT is, therefore, committed to the justificatory force of emotions

themselves, where no inferential steps or separate beliefs are needed to secure justification. Much like perceptions, which reliably lead to true beliefs, the immediate justificatory force of emotions is often taken to be based on their being reliably related to evaluative properties. So, fear is reliably triggered by dangers while anger is reliably triggered by offences, for example. If emotions are reliably correlated to the relevant values, then we might think that they provide defeasible reason for associated evaluative beliefs.

The evaluative judgements we are interested in are those which spell out in conceptual terms what we might think is given non-conceptually in a given emotion. We are, therefore, concerned with evaluative judgements of the form: object x is F, where F is the evaluative property characteristic of each emotion type. Fear of a dog would, for example, provide defeasible justification for the belief 'this dog is dangerous'.

5.1.1 Emotion and Understanding

It is important to note that JT is not the only option available for securing emotions a strong epistemic role. If emotions lack justificatory force, they can still play motivational roles in prompting agents to search for reasons for their emotion. These may be reasons which the agent would not have searched for in the absence of the emotion, and which would then justify relevant evaluative beliefs. Although this motivational role of prompting search for reasons might at first seem to constitute a weaker epistemic role for emotions than the normative role granted by JT, the epistemic goal of each role is actually distinct and the goal of former might be of greater value than the goal of latter.

This is indeed what Brady (2013) argues. For Brady, the epistemic role of emotions is not to justify beliefs but to promote evaluative insight or understanding. The distinction here is between emotions justifying beliefs independently, and emotions promoting understanding about the explanatory relations between beliefs, facts in the world, and ourselves. Understanding, for Brady, is of more value than knowledge, and justified belief, so, in effect Brady (2013: 156) argues that JT *underestimates* the epistemic role of emotions by having as its goal the justification of beliefs rather than evaluative understanding. While evaluative beliefs give us knowledge *that* a certain evaluative property holds of a certain object, 'understanding

emotional objects and events is a matter of understanding why the objects and events have the evaluative properties that they do' (Brady 2013: 142).

I agree that emotions prompt agents to understand why certain evaluative properties obtain, and that doing so often leads to crucial evaluative understanding, as well as self-understanding. If one becomes jealous of an ex-partner for example, this jealousy can motivate one to consider why one feels this way. One can discover repressed romantic feelings for them, or perhaps discover some personal insecurity that would otherwise have gone unacknowledged. In either case, investigating the reasons for one's emotions is a crucial means for gaining self-understanding. In cueing us to what things matter to us, our emotions can cause us to realize what we hold dear, as well prompt us to re-evaluate what *should* matter to us. I will not dispute this crucial role emotions play in promoting evaluative understanding. I will however dispute Brady's (2013) stronger claim that this is the *only* epistemic role emotions play.

I will defend the Justification Thesis, firstly, because it captures our first-hand experience of emotion based epistemic insight. We often form beliefs based on our emotions, without engaging in inquiry about our emotional state. Our emotions often feel to be sufficient reason to believe. Secondly, we might think that the more demanding epistemic achievement of evaluative understanding depends on emotions playing a justificatory role in generating evaluative knowledge in the first place, although I will not argue for this here. Lastly, and most importantly, we are interested in the epistemic role of anger, in particular. We saw that one of the most crucial epistemic roles advocated of anger was the emotion's ability to respond to injustices, and importantly to do so under conditions of structural oppression as an outlaw emotion. Recall that we have as a methodological constraint the commitment not to abstract away from the social and political reality of anger. We should, therefore, take the epistemic potential of outlaw emotion seriously. Now, outlaw emotion cases are likely not going to be cases where evaluative understanding is a feasible epistemic goal. These are cases where the emotion itself feels 'outlaw' to the agent and where the agent typically lacks the resources to uncover the reasons for her emotion. Grimm (2006) writes that:

> Understanding...differs from knowledge in that it involves the ability to give reasons, or grasp causes, or make connections—whereas genuine, reliable knowledge need involve none of these things. (185)

Outlaw emotion cases are likely to be ones where the reasons for one's emotion are not easily available. The epistemic goal of outlaw emotions is, therefore, more likely to be knowledge than understanding. This makes sense seeing as outlaw emotion cases occur in non-ideal circumstances which likely jeopardize the agent's ability to aim for the more demanding epistemic goal of understanding.

Indeed, Brady (2013) says that understanding is the epistemic goal of emotions only 'in those situations where we could become aware of the reasons that our emotions reliably track' (135). The thought is that that:

> we ought to form our evaluative judgements on the basis of genuine reasons where these are available, rather than resting content with our (reliable) emotional responses. (135)

As outlaw emotion cases are unlikely to fulfil this criterion, understanding is unlikely to be their epistemic aim. I don't mean to deny that outlaw emotions can lead to evaluative understanding over time, indeed their role in evaluative understanding is likely indispensable, but they typically do not allow evaluative understanding on their own. Often, interaction with other individuals that share the outlaw emotion, as well as long-term personal development, is necessary, in addition to the outlaw emotion itself, to achieve evaluative understanding. I, therefore, take the Justification Thesis to deserve attention and defence.

5.2   Objections

Maybe JT gets things wrong. Maybe emotions just seem to provide us with reasons for belief, but they are in fact epistemically superfluous. Brady (2010, 2011, 2013) has presented a series of arguments against JT to this effect. Brady argues that emotions do not provide reasons for evaluative judgements, he says 'emotional experiences fail to constitute reason or evidence of any kind for evaluative judgements' (Brady, 2013: 8). For Brady, 'reasons for evaluative beliefs are…considerations that emotional experiences enable us to discover, rather than emotional experiences themselves' (2010: 124). He writes that 'it is our recognitional capacity, rather than our emotional experience, that…justifies our evaluative beliefs' (127), therefore, 'in normal circumstances emotional reactions are epistemically unnecessary' (129).

Brady (2011: 139-40) makes three related claims aimed at motivating scepticism of JT:

> 1. That emotions themselves are reason-responsive.
>
> 2. That when asked why one holds an emotion-based evaluative belief, one typically refers to the reasons for the emotion rather than citing the emotion itself as a reason.
>
> 3. That we do not typically trust our emotional responses as reasons to believe, but take them only to be reasons to search for non-emotional reasons for one's belief.

The thought, therefore, is that emotions cue us to the reasons for which we are experiencing the emotion, but it is these reasons, rather than the emotion, that have justificatory force with respect to evaluative beliefs. The emotion's role is merely to draw the agent's attention to those reasons which justify both the emotion and the associated evaluative belief. Most would grant that emotions cue us to reasons, and often motivate us to search for reasons, the question is whether emotions also have justificatory force of their own.

We have seen that JT is the emotional analogue of an intuitive epistemic thesis held of perceptual experiences. The fact that emotions admit of reasons has been thought to undermine the analogy between emotions and perceptions, thereby undermining JT as a plausible epistemic thesis for the emotional realm (Brady 2010; Cowan 2018)[110]. The disanalogy between emotions and perceptions, with respect to their reason-responsiveness, challenges JT by raising two related concerns. First, immediate justification seems to be a feature of foundational evidence, which in having reasons, emotions are not. JT, therefore, faces the challenge that emotions do not seem to constitute foundational evidence. Second, the reasons that justify the emotions can presumably justify the relevant evaluative belief directly, making the emotion epistemically redundant or superfluous. It seems that in the best case then, emotions provide mediated rather than immediate justification as JT claims, while in the worst-case scenario, emotions have no justificatory force at all. These concerns generate two objections against JT, I will consider each in turn.

---

[110] The analogy between emotion and perception, regarding JT, can be defended in three ways against the charge that emotions have reasons: by arguing that perceptions are actually had for reasons (see Siegel 2017 for a view of perception in line with this), by denying that emotions are had for reasons, or by arguing that JT is not undermined even though emotions are had for reasons. My aim is not to defend an analogy with perception, recall that I am only concerned with how defending that JT holds for emotions. However, given my previous chapter, where I was concerned with providing an account of emotions as reason-responsive, the second option for defending the analogy would not be open to me. The defense of JT I provide in this chapter would support an attempt to defend the analogy of the third type.

5.2.1   Dependency Objection

We have seen that emotions themselves have reasons. There are reasons for sadness and joy, anger and guilt. These reasons are often available to us by non-emotional means, though they need not be. In chapter 4 I proposed an account of emotions as reason-responsive. We can ask ourselves why we have an emotion, whereas why questions do not typically arise of perceptions. This is because although we might have causal reasons for perceptual experiences, most take us not to have normative reasons for perceptual experiences. This disanalogy between emotions and perceptions drives forceful objections against JT. The first I will consider is that JT does not hold of emotions because emotions do not provide foundational evidence for relevant evaluative beliefs (see Cowan 2018). Perceptual experiences provide immediate justification because they are the rock-bottom of the justificatory story:

> Experiences are able to provide justification that is foundational because they lie beyond justification and unjustification. Since they are passively received, they cannot manifest obedience to anything, including rational norms, whether epistemic or otherwise. Since unmotivated by reasons, they can serve as *foundational* sources, as regress-stoppers. (Sosa 2007: 46)

By foundational I will mean the 'rock bottom' of the justificatory story. An experience is foundational then when it admits of no justification itself. As emotions are not the rock bottom of the justificatory story, given that they can themselves be justified, JT is challenged because emotions cannot provide immediate justification to evaluative beliefs. This is the objection. It can be summarized as follows:

P1: There are reasons for emotions.

P2: Only experiences that are not had for reasons can be foundational sources of epistemic justification.

P3: If P1 and P2 then emotions are not foundational sources of epistemic justification.

P4: Only foundational sources of epistemic justification can confer immediate prima facie justification to beliefs based on them.

P5: If P3 and P4 then emotions cannot confer immediate justification to beliefs based on them.

C: Emotions do not provide immediate prima facie justification for evaluative beliefs based on them.

The conclusion is a denial of JT. In chapter 4 I argued for P1, so I will not be challenging it. The objection makes two assumptions that can be resisted, however. First, that only foundational experiences can provide immediate justification (P4) and second, that in having reasons emotions are precluded from being foundational sources of justification (P2).

P4 should come into question once we consider that other sources of justification seem to violate this premise. Memory foundationalism for example is perhaps the most popular theory for the justification of memory-based beliefs (Senor 2014). Memory foundationalism takes memories to provide prima facie immediate justification for memory-based beliefs, despite memories themselves admitting of justification (Schroer 2008). Testimony is another case where an epistemic source that itself admits of justification is often thought to provide immediate prima facie justification the beliefs of a listener (Adler 2017). We therefore have reason to deny P4, given that we tend to think that justification in the realm of memory and testimony can be immediate, despite memory and testimony not constituting rock bottom sources of justification.

Whether or not foundationalist accounts of the epistemology of memory or testimony succeed, we have reason to challenge P2 on independent grounds. P2 says that because emotions admit of justification themselves, they are not foundational sources of justification. This however assumes that emotions merely preserve the information their reasons supply them with. Reasons for emotions either justify emotions directly or they feature in the emotion's 'cognitive base' (perceptions, memories or beliefs) which then justify the emotion[111]. Reasons for emotions can be taken to be non-evaluative features of the world, namely properties of the objects of one's emotion, on which evaluative properties supervene[112]. When afraid of a bear for example, the reasons for your fear are physical properties of the bear (such as sharp teeth and claws) on which the formal object of fear, danger, can supervene. It is your emotion that represents these features of the object of fear as possessing the evaluative property of danger.

---

[111] In theories that postulate cognitive bases for emotions, reasons for emotions feature in the non-evaluative content of perceptions, memories or beliefs, on which the emotion is based (see Deonna and Teroni 2012). When one is angry at one's rival for example, the rival is presented in a perceptual experience, a memory, or he features in a belief. More precisely, it is properties of your rival that feature in your emotion's cognitive base, those properties which make him the target of anger rather than fear, for example. Perceptions, memories and many beliefs are thought to represent non-evaluative properties of the object of your emotion, on which the evaluative properties of emotion types supervene.

[112] This is compatible with the way I construed affective reasons in chapter. see fn.95 in particular.

If emotions have evaluative content, be this intentional or phenomenal content, while their reasons (and cognitive base) are non-evaluative, emotions might be epistemically foundational regarding evaluative properties. In other words, emotions might be the 'rock bottom' of the justificatory story for evaluative beliefs based on them, as emotions are not justified by cognitive bases with evaluative content nor by reasons that themselves feature evaluative properties (see Cowan 2017).

The dependency objection, therefore, does not prove fatal to JT. Both premises P2 and P4 can be challenged. Emotions seem to be able to provide immediate prima facie justification to evaluative beliefs based on them despite themselves admitting of reasons. That being said, emotions might still turn out to be epistemically redundant. This is the objection to which I now turn.

5.2.2 *Superfluity Objection*

There is a problem that relates to the fact that emotions are had for reasons, namely how emotions can act as reasons for evaluative beliefs when the reasons that justify one's emotion are exactly those that could justify the relevant evaluative belief had there been no emotional experience. Brady (2013) writes:

> The fact that the bull in the farmer's field has sharp horns and is advancing rapidly towards me is a good reason for me to be afraid and a good reason for me to judge that the bull is dangerous. Or the fact that Jones keeps changing his story under questioning, refuses to meet his interlocutor's eyes, and stands to gain financially from testifying against the defendant are all good reasons for us to feel mistrust towards Jones and judge that he is untrustworthy as a witness for the prosecution. (113)

Imagine you are angry at your friend for having shared one of your deepest secrets, and this causes you to form the belief that your friend has wronged you. Your emotion is justified if you have reason to think that he indeed shared your secret. JT would have it that your belief that your friend has wronged you, is defeasibly justified by your anger at your friend. But the fact that your friend shared one of your deepest secrets is reason enough to form the belief that your friend has wronged you. It seems, therefore, that your anger at your friend is epistemically superfluous. It seems to be an unnecessary step in a justificatory story. The objection goes as follows:

P1: Emotions have reasons.

P2: Emotion-based evaluative beliefs can be justified by the very same reasons that justify the emotion on which the belief is based.

P3: If an evaluative belief can be justified by the very same reasons that justify the state on which it is based, that state is epistemically superfluous in the justification of the evaluative belief.

P4: If P1, P2 and P3, then emotions are epistemically superfluous in the justification of evaluative beliefs.

P5: If emotions are epistemically superfluous in the justification of evaluative beliefs then they do not provide immediate prima facie justification for evaluative beliefs.

P6: If P4 and P5, then emotions don't provide immediate prima facie justification for evaluative beliefs

---

C: Emotions don't provide immediate prima facie justification for evaluative beliefs

There are a number of ways of responding to this objection. An initial point to make involves the distinction between redundancy and incapacity. The objection implies that emotions are not *required* for the justification of evaluative beliefs. This is compatible, however, with emotions being capable of justifying evaluative beliefs. The beliefs might just be 'justificatorily overdetermined' (Cowan 2016). In other words, the claim that emotions aren't required to justify evaluative beliefs does not imply that emotions lack justificatory force. P3 is challenged as emotions might be capable of conferring defeasible justification, even if we grant that they are unnecessary for the justification of the relevant beliefs. Hence, the conclusion is, in this way, blocked. The Superfluity Objection therefore does not establish the stronger claim that emotions provide *no* reason for evaluative beliefs, only that the emotion is not necessary for the justification of these beliefs.

A similar response to the Superfluity Objection involves highlighting that all the objection establishes is that there are separate ways a relevant evaluative belief can gain prima facie justification. Deonna and Teroni (2012) and Pelser (2014) develop responses of this type and hold JT to remain unthreatened. They take the objection merely to establish that there are two main justificatory routes to a prima facie justified evaluative belief – one emotional and one non-emotional. That a belief be justified via one route does not imply the epistemic superfluity of the alternative justificatory route. That there be a non-emotional justificatory

route to the evaluative belief, then, does not threaten the ability of emotions to confer prima facie justification, hence P3 is challenged, the objection's conclusion is negated and JT stands.

Responses to the Superfluity Objection present in the literature typically target P3. Such responses manage to defend JT but might prove disappointing in doing so. Theorists that seek to afford emotions an important epistemic role might be disappointed by what JT amounts to once the Superfluity Objection has been addressed. As we have seen, the outcome seems to be that emotions have justificatory force, but that they are not required for the justification of evaluative beliefs, nor is there anything particularly special about the epistemic role emotions play.

I take the responses to the Superfluity Objection put forward in the literature to provide adequate defences of JT. In the remainder of this section however, I will outline a response to the objection that seeks to secure emotions a more distinctive epistemic role than existing responses have granted them. My response falls out of arguments made in chapter 4. I will challenge not P3 of the Superfluity Objection, but P2.

The problem with the second premise of the Superfluity Objection is that it assumes that the reason-relation characteristic of emotional justification is the same as the epistemic reason-relation. I argued against this assumption in Chapter 4, where I proposed that we take reasons for emotions (which I called affective reasons) to constitute a sui generis basic reason-relation. I argued that we have reason to replace the traditional dichotomy between epistemic and practical reasons by a trichotomy that makes room for affective reasons. In chapter 4 (sec 4.4) I spelt out how emotional reason-responsiveness differs from the reason-responsiveness of beliefs. Specifically, I argued that affective reasons are characterized by strictness and sufficiency. Sufficiency accounts for the fact that one affective reason is sufficient to justify an emotion, while strictness captures the fact that there can be reason to feel sad even when there is reason to feel happy. In such cases there is reason to feel both happy and sad or a mixture of both. The affective reasons do not cancel each other out. Beliefs on the other hand are supported by weighing up separate reasons, none of which are necessarily sufficient to justify a belief, and which can cancel each other out.

If emotions and beliefs bear distinctive types of normative relation to the facts that support them, as I have argued, there will be cases in which the reasons for one's emotion are not

reasons that could justify the relevant evaluative belief. This would deny P2, or rather, significantly restrict the cases in which P2 holds. Consider the following example: a non-emotional case of (un)trustworthiness. If one has reasons to believe in Maria's trustworthiness, as well as reasons to doubt it, one weighs the evidence for and against Maria's trustworthiness in the formation of the relevant belief. One of the following scenarios occurs:

> 1. Evidence[113] in favour of Maria's untrustworthiness outweighs evidence against it, providing prima facie justification for the belief that Maria is untrustworthy.

> 2. The evidence doesn't tip in favour of either Maria's trustworthiness or her untrustworthiness. There is not enough evidence to support a prima facie justified belief either way[114].

> 3. Evidence in favour of Maria's trustworthiness outweighs evidence against it, providing prima facie justification for the belief that Maria is trustworthy.

Now let's look at the emotional versions of these scenarios. In all scenarios one presumably has reason to feel suspicion, there is, after all, some evidence of Maria's untrustworthiness in all three scenarios. Given the strictness and sufficiency of affective reasons, an independent reason for feeling suspicion is enough to justify the affective attitude, even if there are reasons for distinct affective attitudes. By responding to a reason for *feeling* suspicious of Maria, you experience suspicion. This feeling of suspicion provides immediate prima facie justification to the evaluative belief that Maria is untrustworthy, according to JT[115]. This will lead to variants of the above three scenarios where all of them involve a prima facie justified evaluative belief regarding Maria's untrustworthiness. The three scenarios become:

> 1*. An evaluative belief that Maria is untrustworthy is justificatorily overdetermined, or there are two routes to the justification of the evaluative belief. The evaluative

---

[113] I am using 'evidence' and 'epistemic reason' interchangeably.

[114] When a belief is ungrounded it is common to take it to lack even prima facie justification (see Senor 1996)

[115] I'm assuming JT for the sake of constructing intuitive example scenarios. This is not begging the question against the Superfluity Objection because I am assuming JT only to bring out the difference in normative support between emotions and the normative support of evaluative beliefs that are not based on emotions. It is these differences in normative support that provide defense against the Superfluity Objection, and these differences do not depend on JT (see chapter 4 sec 4.4).

belief regarding Maria's untrustworthiness is prima facie justified both/either by the feeling of suspicion and by the reasons for feeling suspicion.

2*. An evaluative belief that Maria is untrustworthy is prima facie justified by the feeling of suspicion. The facts that justify the feeling of suspicion are not sufficient to provide prima facie justification for the evaluative belief directly (as seen in scenario 2).

3*. An evaluative belief that Maria is untrustworthy is prima facie justified by the feeling of suspicion. The facts that justify the feeling of suspicion cannot provide direct prima facie justification for the evaluative belief for they are outweighed by further epistemic reasons (as seen in scenario 3).

In non-emotional scenarios 2 and 3, one does not form a prima facie justified belief that Maria is untrustworthy. In scenario 2* and 3* it is the emotion that provides prima facie justification for the evaluative belief that Maria is untrustworthy. What's going on in these cases? In scenario 2*, the reasons that count in favour of Maria's untrustworthiness, count in favour of one's emotion in a different way to how they count in favour of the belief that Maria is untrustworthy. Namely, these reasons are sufficient for the emotion and not for the belief.

To focus solely on the dynamics of epistemic reasons, let's briefly concentrate on non-emotional scenario 2. In this scenario, there are two main ways in which there can be insufficient epistemic reason for the belief that Maria is untrustworthy. Perhaps there simply is not sufficient epistemic reason in favour of Maria's untrustworthiness. Maybe there is only one such reason. Alternatively, perhaps there would have been sufficient epistemic reason for the belief that Maria is untrustworthy, had there not been further epistemic reasons at play that cancelled them out. In the former case, the sufficiency of affective reasons will account for the different epistemic outcome in 2* as compared to 2. We saw that affective reasons are characterized by sufficiency, whereas epistemic reasons are not. This means that, whereas affective reasons are sufficient for the justification of emotions, epistemic reasons need not be. In 2*, that there is reason for one's emotion is sufficient to justify it. The evaluative belief that Maria is untrustworthy, therefore, is based on one's emotion, given that there is insufficient epistemic reason to support this belief directly.

Alternatively, if there would have been sufficient reason for the evaluative belief in scenario 2, but further considerations cancelled them out, then the strictness of affective reasons is also relevant to the epistemic difference between scenario 2 and 2*. The further considerations that cancel out reasons for belief in Maria's untrustworthiness in scenario 2 can be evaluative or non-evaluative epistemic reasons. Recall that affective reasons are considerations that count in favour of evaluations. Considerations that count in favour of evaluations can also provide normative support for beliefs and actions. Indeed, we saw that a given fact can act as a reason for distinct types of attitudes, be they epistemic, affective, or practical. Given the strictness of affective reasons, they cannot be cancelled out by other affective reasons. They, therefore, cannot be cancelled out by considerations that count in favour of evaluations, or what I called 'evaluative reasons' for short. This means that reasons to feel grateful towards Maria, for example, cannot cancel out reasons for suspicion. Epistemic reason to believe Maria to be a generous person, however, might outweigh reasons to believe Maria is untrustworthy. In scenario 2 one might also have non-evaluative reasons against Maria's untrustworthiness. These will be any considerations that count in favour of her being trustworthy, but that aren't evaluative. For example, that Maria is a punctual person might be a reason in favour of her trustworthiness. These non-evaluative reasons can cancel out evaluative epistemic reasons in scenario 2, such that there is insufficient reason for the belief that Maria is untrustworthy. The non-evaluative reasons cannot, however, cancel out affective reasons for suspicion in scenario 2* because, in demonstrating sufficiency, only *sufficient* non-evaluative epistemic reasons will be able to do so. As scenario 2 does not involve sufficient epistemic reason for belief either way, scenario 2* is one where there is not sufficient epistemic reason to cancel out affective reasons for suspicion.

Scenario 2* involves a prima facie justified evaluative belief that could not have been justified independently of the emotion. Emotions therefore seem to play important epistemic roles in cases where there is insufficient reason for evaluative beliefs directly. In these cases, the emotion is the only means for ensuring the prima facie justification of relevant evaluative beliefs. P2, therefore, does not seem to hold of these scenarios. What about scenario 3* and compared to 3? In scenario 3* the reasons that could justify the evaluative belief directly are outweighed by other epistemic reasons. The very same reasons are, however, not necessarily outweighed in their normative support of the emotion, seeing as affective reasons are strict and demonstrate sufficiency. We will see that scenario 3*, therefore, involves a prima facie justified evaluative belief that could not have been justified independently of the emotion.

The only way the emotional version of scenario 3 does not generate a prima facie justified belief based on one's suspicion, is if one's affective reasons are outweighed by epistemic reasons. For the affective reasons to be outweighed by the epistemic reasons against Maria's untrustworthiness, however, there needs to be sufficient non-evaluative epistemic reason for Maria's trustworthiness. This is because affective reasons demonstrate strictness and sufficiency. In being strict, affective reasons will not be cancelled out by evaluative epistemic reasons, while in demonstrating sufficiency, affective reasons require sufficient epistemic reason against them so as to be cancelled out.

The epistemic reasons for Maria's trustworthiness that outweigh reasons against it, in scenario 3, might be a mixture of evaluative and non-evaluative considerations. Affective reasons will be immune to being outweighed by other affective reasons, given strictness. Affective reasons are considerations that count in favour of evaluations. We saw that considerations that count in favour of evaluations can also provide normative support for beliefs and actions, however. Strictness, ensures that affective reasons are not outweighed by these evaluative considerations. Therefore, although in the epistemic case there will be sufficient (evaluative and non-evaluative) epistemic reasons to outweigh reasons to believe in Maria's untrustworthiness, there might not be sufficient purely non-evaluative epistemic reason to do so. Given their strictness, affective reasons remain immune to considerations that in the epistemic realm might play a determinate role in ensuring that one has sufficient reason to outweigh conflicting evidence. If there is not sufficient epistemic reason against Maria's untrustworthiness, once evaluative epistemic reasons have been discounted, then one's epistemic reasons do not cancel out one's affective reasons. This means P2 is not true in many cases where there is sufficient reason to form a belief that goes against one's emotion.

Scenario 3 involved a prima facie justified belief that Maria is trustworthy (supported by epistemic reasons), while I have said that 3* involves a prima facie justified belief that Maria is untrustworthy (supported by the feeling of suspicion). If 3* is scenario 3, just with the feeling of suspicion added, then scenario 3* actually involves the agent holding two prima facie justified beliefs: that Maria is trustworthy and that Maria is not trustworthy. This is not an ideal epistemic state to be in, but it is nonetheless one where the emotion in 3* is responding to reasons that were present in scenario 3, and which would otherwise not be playing a justificatory epistemic role. Perhaps in 3* one suspends judgment on whether Maria

is trustworthy. This still amounts to a distinct epistemic outcome as compared to scenario 3, so the emotion is still playing a distinct epistemic role.

Scenarios such as 3* are one's that might be called cases of rational recalcitrance. One has a belief that Maria is untrustworthy, justified by epistemic reasons, while still experiencing an emotion of suspicion for which there is reason. Outlaw emotion cases seem to be cases of this type. I will say more about outlaw emotions below, for now I mean only to highlight that such cases are structurally similar to scenarios in which emotions seem to play an epistemically indispensable role. Indeed, outlaw emotion cases have actually been mentioned by proponents of JT, in response to the Superfluity Objection, as ones where the emotional route is the only justificatory route available (Deonna and Teroni 2012b; Tappolet 2014). This fits well with the claims made by feminist philosophers about the epistemic value of outlaw emotions.

The Superfluity Objection, therefore, applies only to situations similar to scenario 1*, where there are epistemic reasons that can provide prima facie justification to the evaluative belief that Maria is untrustworthy directly. The objection, therefore, only applies to a subset of cases, rather than holding true of all, or even typical, instances of emotion epistemology[116]. And, of course, we must remind ourselves that even in cases where the objection does apply, it does not undermine JT, as we have seen from responses to the Superfluity Objection present in the literature, that target P3. The objection only manages to establish that that the emotion is not epistemically necessary. I have argued that the objection's success is even more limited: it only manages to establish that emotions are epistemically unnecessary in a subset of scenarios. Therefore, emotions are at least sometimes, and plausibly often, necessary for the justification of evaluative beliefs.

---

[116] Note however that the emotion, following JT, provides immediate prima-facie justification, while in scenario 1* the epistemic reasons can only provide mediated or inferential prima facie justification to the belief. It is unlikely that any individual epistemic reason gives one immediate prima facie justification for the belief that Maria is untrustworthy. Even if some pieces of evidence are perceptual (such as you having seen Maria steal an ice cream cone for example), perceptual experiences only provide immediate prima facie justification for empirical beliefs (that Maria has stolen the ice cream cone), their role in support of the evaluative belief (that Maria is untrustworthy) would therefore not be that of immediate prima facie justification. Even in a number of scenarios in which the superfluity objection arises then, emotions may remain the only source of *immediate* justification.

Recall Brady's (2011: 139-40) three related claims aimed at motivating scepticism of JT.

> 1. That emotions themselves are reason-responsive.
> 2. That when asked why one holds an emotion-based evaluative belief, one typically refers to the reasons for the emotion rather than citing the emotion itself as a reason.
> 3. That we do not typically trust our emotional responses as reasons to believe, but take them only to be reasons to search for non-emotional reasons for one's belief.

The first observation generated the two objections that I defended JT against. Observations 2 and 3, on the other hand, are first personal observations about the psychology of emotional epistemology that are meant to motivate the force of these objections. We need not think that first personal experience tracks epistemic facts, however. I have defended JT against the two epistemic objections that these first personal observations motivate. I have, therefore, provided argument against their epistemic upshots. To conclude my response to Brady, however, it is worth considering whether and why his phenomenological observations might hold true.

I take observation 2 to be a far more accurate observation than 3. Observation 3 seems to simply get things wrong as a claim about *typical* emotional experience. After all, the complaint that emotions are irrational and tend to mislead us is precisely motivated by the fact that we tend to form beliefs based on them, without searching for reasons for our emotions. There is empirical evidence for the propensity to form beliefs based on feelings (Clore and Gasper 2000). In addition to this, I think our phenomenological experience of emotions often involves feeling the emotion to be apt[117]. If this is the case, then emotions do not typically cause or give us reason to search for non-emotional reasons for beliefs.

What about observation 2? If, that emotions are reason-responsive, does not threaten their capacity to provide prima facie justification to evaluative beliefs, why do we typically answer why questions regarding evaluative-beliefs based on them by reference to the reasons for our emotions, rather than by reference to our emotions? It's important to note that we often cite

---

[117] This fits the story I told of attention being drawn to reasons for emotions in emotional experience, in chapter 4. Furthermore, the idea that emotions present their objects as meriting the emotional response is not uncommon in the philosophy of emotion literature (see Mitchell 2017).

evidence beyond our own experience in response to why questions in non-emotional cases as well. If asked a why question about a belief based on testimony, you often cite the reasons for taking the testimony to true, rather than citing the testimony itself[118]. Similarly, when asked why you believe it is raining, you might point out the window (Pelser 2014). Citing reasons for one's mental states, in response to why-questions regarding beliefs based on them might, however, be more prevalent in the emotional realm for a few reasons. First of all, emotions hold a denigrated epistemic status in many social contexts[119]. This means that one will be more inclined to cite non-emotional reasons in response to why-questions. Brady's observation 3 then, may be false in societies where emotions do not hold a denigrated epistemic status. It may even be false of agents whose emotional responses have come to be considered extremely reliable. Secondly, emotions might admit of defeaters at more levels than typical sense-perceptions, as they are often based on other mental states such as sense-perceptions, beliefs or memories. The justificatory force of emotions can be defeated by evidence that speaks against both the emotion, as well as the cognitive base of the emotion. This might explain why non-emotional reasons are cited in support of emotion-based beliefs more often than in perceptual cases[120]. A final reason to avoid mentioning emotions in response to why questions, is that they are often seen as unreliable. I turn to this object now.

5.2.4 Objection from Unreliability

Unreliability is a common charge made against emotions (Goldie 2004a; Pelser 2014; Tappolet 2018c). The force of this charge is however, exaggerated. First let's consider what the charge actually amounts to. If the claim is that emotions are unreliable at generating ultima facie justified evaluative beliefs this doesn't threaten JT, for JT is a claim about prima facie justification. If on the other hand the unreliability charge is stronger and holds that

---

[118] If asked why you believe enzymes catalyze biochemical reactions, the response 'because my biology teacher told me' is not going to be very satisfying. Typically, you will cite reasons for the truth of the belief itself, such as the functional properties of enzymes, rather than citing its epistemic basis, namely the testimony, as a reason for your belief.

[119] That we do not cite emotions as the basis of beliefs might be evidence that something akin to what Fricker (2007) has termed a hermeneutical injustice, is at play. Such injustices occur when shared conceptual resources are lacking to make sense of experience. In the emotion case we might think that our society is hostile to the development of a conception and practice of emotional knowledge. If such a story is true, then it is no wonder we avoid referring to our emotions whenever attempting to communicate why a certain evaluative property is instantiated, we lack the hermeneutical resources to do otherwise.

[120] Mitchell (2017) makes a similar point.

emotions, in being unreliably related to true beliefs, do not even provide prima facie justification, then JT is threatened.

Should we hold such a strong view however? I don't think we have reason to. The evidence levied in favour of the unreliability of emotions I think falls into one of the following categories. A) Emotions produce false positives. (Ex: 'You gave me a fright, I thought you were a burglar!') B) Emotions prompt actions we later regret (Ex: 'Sorry I hit you but I thought you were a burglar!') C) Emotions skew our epistemic landscape[121]. (Ex: 'Now that you've frightened me I keep being distracted by noises outside and thinking it could be a burglar.')

That emotions generate false positives is compatible with emotions reliably responding to reasons. That emotions are triggered in an excess of cases does not undermine their reliably tracking cases in which they are called for. This means that emotions might be reliable enough to provide evaluative beliefs prima facie justification, after all, prima facie justification is a weak epistemic achievement. The view that emotions generate false positives is particularly plausible on a view where emotions are equated to affect programs, however. We saw that affect programs are fast, modular and serve an evolutionarily hardwired function. Indeed, why should an emotion that has a basic evolutionary function of fight or flight, for example, be a reliable process for generating true beliefs in the modern day where we do not have the same concerns or face the same threats? In chapter 2 I argued against the basic theory of emotion, that equates emotions to affect programs of this sort. Emotions emerged as heterogeneous phenomena that have evolved, biologically and culturally, to play far more complex and less modular functions. The evolutionary argument in favour of the unreliability of emotions is, therefore, based on a mistaken understanding of emotions as evolved phenomena[122].

---

[121] See Goldie (2004a)

[122] What about when emotions have culturally evolved to get things wrong? Some people feel disgust towards expressions of gender and sexuality that transgress cultural norms, for example. Against what I have implied then, the cultural evolution of emotions doesn't speak to their reliability at tracking evaluative properties, indeed it might threaten it. I agree that emotions can be molded, on multiple time-scales, to get things wrong, especially under oppressive conditions. That emotions can get things wrong is not enough to establish their unreliability. After all, these emotions are still getting things right relative to their oppressive context, they are appropriate relative to that community. This is more than would be allowed on a simplistic understanding of emotions as reflex affect programs. I mean only to cast doubt on the support that simplistic evolutionary stories might provide the claim that emotions are unreliable. For if the misguided evolutionary story were true, then we would have reason to doubt emotions to be reliable at tracking evaluative properties in modern day life.

However, we in fact need not depend on arguments made in previous chapters to defend emotional reliability. For, in so far as a simplistic evolutionary view explains the prevalence of false positives (given that it would be more adaptive to be safe than sorry), such an evolutionary story does not predict of emotions that they generate so many false positives as to be completely unreliable. Constantly generating false-positives would not have been adaptive. We, therefore, have reason to think that, when an emotion occurs, it is more likely that there is reason for the emotion, than there would be had no emotion occurred. This basic sense of reliability is hard to deny of emotions and should suffice for prima facie justification.

That being said, we might have reason to think emotions are more reliable than this, as much of the evidence in favour of the epistemic unreliability of emotions misses the mark. Situations of type B) above do not actually speak to the epistemic unreliability of emotions. I included B) because I think that the putative evidence offered in support of the unreliability of emotions often stems from a failure to distinguish between the emotion's practical effects as opposed to its normative rationality. In many cases the agent feels they would have been better off at preserving their interests had they not experienced and acted upon the emotion. This might be true in given circumstances, but it doesn't make the emotion unreliable at tracking reasons.

Similarly, evidence of type C) involves emotions having negative epistemic effects, but these effects need not include being unreliable at generating true beliefs about the object of the emotion. The negative epistemic effects caused by emotions might be seen as violations of norms of rationality, outlined in chapter 4, as they might involve violating rules of inference, for example. Evidence of type C) is that for which there is most empirical evidence. Emotions have been observed to bias downstream effects on one's reasoning, including the validity of one's judgements of risk, benefit and competence (Finucane et al. 2000). That emotions can have these effects, however, does not undermine their epistemic reliability. In feeling fear of rock climbing, or pride in one's driving abilities, one might systematically overestimate the risk of rock climbing, as well as overestimate one's competence at driving, but these effects do not undermine the reliability of your fear or pride in producing true beliefs regarding the object of your emotion. That is to say, your fear and your pride might be justified in cases where they have deleterious epistemic effects.

Furthermore, errors in risk, benefit, and competence estimates can occur without the agent experiencing an emotion. The agent's beliefs (that they've never been in a car accident, and that they know someone who became seriously injured from rock climbing for example), can bias their reasoning. Indeed, even perceptual priming has biasing effects on our behaviour and reasoning (Ortells et al. 2006). We therefore have reason to think that differences in skewing effects between emotions and other mental phenomena have been exaggerated. Even if emotions were shown to be particularly noxious to reasoning however, we have seen that this would not establish that they are unreliable at generating true evaluative beliefs about their objects. Other well-known empirical evidence has been gathered in support of the indispensability of emotions to proper reasoning (Damasio 2005). Given that some emotional contribution is likely to necessary for the proper functioning of most cognitive processing, emotions might be far more epistemically beneficial than previously thought. Even if this turned out not to be the case, however, a basic sense of reliability is hard to deny emotions.

5.3   Beyond Prima Facie Justification.

I have defended JT against objections present in the literature[123]. In the second part of this chapter, I turn to characterizing an unacknowledged novel objection. I then argue that JT can survive the objection only if appeal is made to insights from social epistemology.

The objections raised of JT in the literature typically concern the prima facie justificatory role of emotions, while neglecting JT's commitment that emotions are meant to supply ultima

---

[123] There is one objection made in the literature that I have not addressed: the Self-Justification Objection (Brady 2011; Cowan 2016; Mitchell 2017). Briefly, this is the objection that if emotions are a source of normative support for evaluative beliefs, then they are also a source of normative support for themselves. As we should not endorse a view that allows emotions to be self-justifying, the objection goes, JT should be abandoned. Brady's (2011) version of this objection presupposes a response-dependent theory of evaluative properties, whereby emotion E self-justifies because it features evaluative content of the sort 'that there are reasons to feel E', and therefore, if emotions are reasons, in being felt the emotion will be a reason for itself (because the emotion involves content regarding its own fittingness). Direct responses to Brady's self-justification objection have been provided by Cowan (2016) and Mitchell (2017). The thought that emotions can be self-justifying can be cashed in a way that does not depend on a response-dependent theory of evaluative properties, however, as follows: The evaluative belief justified by the emotion (that object x has F evaluative property) can itself act as the normative base for the emotion, making the emotion justified by the very belief it provides justification for (Cowan 2017). It is important to note that this objection can arise for any reason-responsive experience that we think provides immediate justification (ex: memories). We therefore have reason to either consider the self-justification involved benign or to think that it might be a feature of reason-responsive states that generate normative support for beliefs, that they cannot confer normative support onto themselves (See Cowan 2017).

facie justification to the relevant beliefs in the absence of defeaters. Prima facie justification, as we have seen, is usually cashed out in terms of defeasibility, such that for a belief to be ultima facie justified, the experience on which it was based must be undefeated. JT is therefore committed to a popular view in epistemology often called defeatism, whereby a defeater clause is included as a condition on knowledge or justification (Bergmann 1997). Proponents of JT endorse a defeater clause, given that that take emotions to provide defeasible justification, even if they typically do so only implicitly.

This commitment can be summarized as follows:

Epistemic Defeatism (ED): Emotions provide ultima facie justification for associated evaluative beliefs provided there are no defeaters at play.

Some perceptual theorists have explicitly endorsed the defeater clause ED. Tappolet (2016) for example writes that:

Your belief that x is fearsome is justified sans phrase if
    (a)   you feel fear towards x,
    (b)   your belief is based on your fear, and
    (c)   you have no reason to believe that your fear is inappropriate. (170)[124]

Requirement 'c' is a negative requirement that seems equivalent to ED. Similarly, Pelser (2014), in a paper defending JT, writes that:

There might be and often are overriding factors that prevent the prima facie justification of a particular source of belief from becoming ultima facie justification without rendering the source impotent to justify in every case. (116)

Unlike objections raised of JT in the literature, I raise a novel objection to the view regarding its commitment to a defeater clause for the ultima facie justification of emotion-based beliefs. In short, ED risks ruling out outlaw emotion cases, which are one's highlighted by feminist philosophers as of paramount epistemic value. These are also cases that endorsers of JT have explicitly taken their view to be able to accommodate (See Tappolet 2014; Deonna and Teroni, 2012: 122).

---

[124] Following Tappolet, I use the term 'sans phrase' and 'ultima facie' interchangeably.

What exactly counts as a defeater under ED? Despite widespread, albeit often implicit, endorsement of ED, the nature and functioning of defeaters is not a topic of much discussion in the philosophy of emotion literature. Seeing as JT is modeled on perceptual epistemology however, we can pinpoint some clear defeaters. We saw earlier that factors that affect the processing of the domain in question are typical defeaters. Being drunk or colour blind are two defeaters for trusting how things perceptually seem to you at a given time. These are typically called undercutting defeaters (Pollock 1986: 39) because they are reasons to doubt the truth of the grounds of one's belief.

Similarly, your belief that a dog is dangerous would be defeated by knowledge that you had been injected with a hefty dose of adrenaline prior to encountering the dog. As emotions are often thought to involve cognitive bases however, we might think that beliefs-based on them have more types of defeaters. Emotions can be defeated by the sort of defeater that interacts with emotional processing systems, such as brain stimulation of the limbic area or having chemically excited one's nervous system, but emotions can also be defeated by those things that defeat their cognitive basis. If you have taken magic mushrooms and become afraid of a dog, your belief that the dog is dangerous is potentially defeated by the drugs acting as a perceptual defeater against your perception of the dog. The justification of beliefs based on emotions with memories or thoughts as their cognitive base would likewise be defeated by defeaters that cast doubt on the functioning of one's memory or one's cognitive capacities for example. Emotion-based beliefs are therefore are likely to admit of undercutting defeaters of two sorts: those that undercut the emotional experience and those that undercut the potential cognitive bases of emotions.

Another plausible defeater is a belief that conflicts with the content of your experience. In the perceptual realm, when your perception of the Müller-Lyer illusion shows the two lines as different in length, for example, your belief that the two lines are in fact the same length acts as a defeater against the justificatory force your perceptual experience. You are therefore not justified in believing that the two lines are of equal length, despite them appearing so. This is an example of a rebutting defeater. These are defeaters which give one reason to hold the negation of the defeated belief (in this case the empirical belief), or to hold a belief that is inconsistent with the defeated belief (Pollock 1986: 38). Similarly, in the emotional realm, a justified belief that the dog in front of you is harmless defeats the justification of any belief to the contrary based on your fear. Pelser (2014), a proponent of JT, writes that:

A particular emotional experience might fail to provide justification for a belief formed on the basis of that experience in case the subject has evidence that counts against trusting her emotional experience or conflicts with the propositional content of the experience. (16)

I, therefore, take emotion-based evaluative beliefs, much like empirical beliefs, to admit of both undercutting and rebutting defeaters. Endorsing ED then, seems to commit us to the view that emotions cannot justify beliefs that conflict with already held beliefs, because the conflicting belief acts as a rebutting defeater against the ultima facie justification of such emotion-based beliefs. We will see that this risks ruling out outlaw emotion cases as they involve conflicting beliefs, that much like the perceptual case, seem to act as defeaters against the justificatory force of the outlaw emotion.

5.4   Outlaw Emotions

Jaggar (1989: 166) writes that under conditions of systemic gender oppression 'we are all likely to develop an emotional constitution that is quite inappropriate for feminism'. People however, 'do not always experience the conventionally acceptable emotions'. Indeed:

> They may feel satisfaction rather than embarrassment when their leaders make fools of themselves. They may feel resentment rather than gratitude for welfare payments and hand-me-downs. They may be attracted to forbidden modes of sexual expression. They may feel revulsion for socially sanctioned ways of treating children or animals. In other words, the hegemony that our society exercises over people's emotional constitution is not total. (166)

Jaggar (1989) writes that it is more likely for members of subordinated groups to experience conventionally unacceptable emotions:

> People who experience conventionally unacceptable, or what I call 'outlaw' emotions often are subordinated individuals who pay a disproportionately high price for maintaining the status quo. The social situation of such people makes them unable to experience the conventionally prescribed emotions: for instance, people of color are more likely to experience anger than amusement when a racist joke is recounted, and women subjected to male sexual banter are less likely to be flattered than uncomfortable or even afraid. (166)

There is a question regarding whether emotions are more encapsulated, so to speak, from dominant ideology than a subject's belief system. Jaggar seems to suggest so. Answering this question in the affirmative, however, will be hard to square with ample evidence on the

emotionally driven implicit biases of people who actually hold explicit egalitarian beliefs. We need not endorse an implausibly strong version of the claim, however. We need only hold that sometimes emotions better track reasons than our belief systems do. To deny this weaker claim would be to deny the genealogy of liberatory movements, and the role of consciousness raising, where collective efforts to make sense of shared emotional experiences led to critical progress (see Fricker 1991, for example).

The view that emotions play a strong epistemic role in generating knowledge is widespread in the feminist literature (Frye 1983; Narayan 1988; Jagger 1989; Bell 2009). One of the most crucial epistemic roles for emotions arises in outlaw emotion cases. This is where emotions pick up on reasons the subject is otherwise insensitive to given the internalization of oppressive ideology. Jagger (1989) coined the term 'outlaw emotions' to refer to these emotions that are 'distinguished by their incompatibility with the dominant perceptions and values' (166). Recall Friedman's (1986) oft cited passage, which illustrates the epistemic value of outlaw emotions:

> A woman who has been taught that a 'woman's place is in the home' may be driven to question this maxim precisely in light of her persistent dissatisfactions and repeated urges to flee from the responsibilities and limitations which structure her domestic life… If her highest principles themselves also include notions of 'appropriate' sex roles, duties to others and the importance of self-sacrifice as an ideal of femininity, then there is not much available among her highest principles to afford an independent standpoint for assessing the maxim about woman's place. Her frustration, grief, and depression, and the motivations to change her life which spring from these sources, may be her only reliable guides. (31)

What exactly is the epistemic role played by outlaw emotions, however? In Friedman's passage it seems that the outlaw emotions are 'persistent' and 'repeated' urges that motivate one to *question* one's oppressive beliefs. Outlaw emotions seem here to be motivating the agent to *assess* their current beliefs about a woman's place in the home. This is best seen as a motivational as opposed to normative, or justificatory, epistemic role. This motivational role is crucial as it is often the fuel necessary for developing a critical stance on one's social position. Outlaw emotions motivate the search for reasons, and are, therefore, likely to be crucial in the pursuit of evaluative understanding under conditions of oppression. Much like with other emotions however, the epistemic role of outlaw emotions is not limited to the motivational. Outlaw emotions have justificatory force and often justify evaluative beliefs directly. Jaggar (1989) says that:

> The most obvious way in which feminist and other outlaw emotions can help in developing alternatives to prevailing conceptions of reality is by motivating new investigations. (167)

But:

> As well as motivating critical research, outlaw emotions may also enable us to *perceive* the world differently from its portrayal in conventional descriptions. (emphasis my own, 167)

Here we see that the motivational role is not the only one proposed of outlaw emotions. Outlaw emotions allow us to perceive their objects in unconventional ways. When Friedman's housewife experiences anger, she sees her situation as unjust. In addition to motivating inquiry then, this anger may provide immediate justification to an evaluative belief that her situation is unfair. Consider the following structurally identical example. Raquel is a woman living under conditions of gender oppression who is on a night out with a group of girlfriends. She feels someone squeeze part of her body and turns to see a man she has never met before. The man smiles at Raquel in acknowledgement of having been the person to touch her, and walks away. Some of Raquel's friends feel excited by what just happened, and encourage her to go talk to the man. Others aren't excited but jealous for not having been the one squeezed. Raquel believes she should be flattered by the attention she has received, as well as proud to have been the one approached, and indeed she does feel a mixture of these emotions, but she also feels uneasy and angry. Raquel forms the belief that what the man did was not acceptable, based on her negative emotional responses, despite it going against her wider set of beliefs and emotional responses, as well as those of her peers.

Raquel's negative emotional response might motivate her to reflect on the event and deliberate about why she responded this way. But in becoming angry, Raquel is perceiving the man's actions as problematic, even if she doesn't understand why. The event strikes her as wrong, while it strikes her friends as flattering and desirable. They perceive the event in different ways given their distinct emotional responses. That the event affectively strikes Raquel as wrong is sufficient to provide prima facie justification to the belief that she has suffered an offence, according to JT. When we experience an outlaw emotion then, it is not only epistemically valuable for motivating inquiry, but because it has immediate justificatory force that allows one to form prima facie justified beliefs about the situation.

This immediate justificatory role is important not only because I take it to match the phenomenology of agents living under conditions of oppression, but also because often the motivational role of outlaw emotions in searching for reasons will not be a straightforward one. Raquel, might not be aware of why she feels angry, nor might it be easy for her to become aware of reasons for her negative emotional response, after all she also feels somewhat proud and flattered, and her friends show no signs of thinking any type of negative response would be appropriate. Jaggar (1989) writes that:

> When unconventional emotional responses are experienced by isolated individuals, those concerned may be confused, unable to name their experience; they may even doubt their own sanity. Women may come to believe that they are 'emotionally disturbed' and that the embarrassment or fear aroused in them by male sexual innuendo is prudery or paranoia. When certain emotions are shared or validated by others, however, the basis exists for forming a subculture defined by perceptions, norms, and values that systematically oppose the prevailing perceptions, norms, and values. (166)

The route from outlaw emotions to evaluative understanding, or even to the mere uncovering of reasons for the emotion, will therefore often be a difficult one. It will be a route that often depends on interaction with others who share one's outlaw emotions, and is likely to be psychologically demanding. In being felt however, outlaw emotions also play an immediate justificatory role in support of evaluative beliefs based on them. Indeed, these evaluative beliefs might be important stepping stones to the success of the motivational role outlaw emotions play. And in so far as the motivational role might take some time to deliver epistemic fruits, the direct normative role entitles the agent to an immediate evaluative belief that is often their first radical epistemic achievement[125].

I have claimed that outlaw emotions, much like emotions in general can act as unmediated justification for associated evaluative beliefs. JT therefore seems to be ideally suited to make sense of this epistemic role of outlaw emotions. In line with JT, the outlaw emotion provides immediate prima facie justification for an evaluative belief. A woman that has been taught that her place is in the home, yet is dissatisfied, is prima facie justified in believing that their

---

[125] The motivational epistemic role of outlaw emotions is likely to eventually lead to evaluative understanding of one's oppression and its features. Beliefs of the sort 'this is sexual harassment' are likely only possible for Raquel to hold as a downstream effect of the motivational role of outlaw emotions. This is unsurprising given that Raquel currently lacks the concept of sexual harassment, as well as lacking full awareness of the reality of the term's referent. Raquel's outlaw emotion will typically provide immediate prima facie justification to beliefs of the sort 'this is wrong' or 'something isn't right about this'. The justificatory route therefore allows quite imprecise beliefs to be prima facie justified, but the importance of these beliefs shouldn't be underestimated, as they are often the agent's first epistemic achievement in resisting their oppressive conditions.

situation is unfair. JT does not require the agent to be aware of one's reasons for the emotion, much less understand them. This is a benefit as outlaw emotions are characterized as spurs to novel insights, so they mustn't presuppose the very knowledge they are meant to shed light on. JT, therefore, seems ideally equipped to account for outlaw emotion cases[126]. We will see, however, that JT actually risks ruling out the epistemic role of outlaw emotions. This is because JT involves the defeater clause ED and outlaw emotion cases are ones we have good reason to think always involve defeaters.

5.5   The Objection from Outlaw Emotions

The objection to JT is generated by taking the epistemic role of outlaw emotions to be a feature that a successful epistemological account of the emotions should capture. Recall that we have as a methodological constraint the commitment not to abstract away from the social and political reality of anger. Outlaw emotion cases of anger are ones that arise under conditions of oppression and that have been advocated by feminist philosophers to constitute one of the emotion's most important epistemic roles. We should, therefore, take outlaw emotion cases to be ones where emotions are of potentially supreme epistemic value. Given this methodological constraint, accounting for the epistemic role of outlaw emotions emerges as a desideratum for an adequate epistemology of emotions.

In line with this, proponents of JT have claimed that JT's ability to accommodate for outlaw emotion cases speaks to the strength and plausibility of their view (see Deonna and Teroni 2012a: 122; Tappolet, 2014). These thinkers take JT to account for the epistemic role of outlaw emotions under conditions of gender oppression (Tappolet 2014), as well as cases of outlaw emotions under conditions of racial oppression (Deonna and Teroni, 2012b). In these cases, proponents of JT grant that the outlaw emotion plays a justificatory role despite the agent lacking awareness of their reasons as reasons, or any intermediary belief about the status of one's emotion or one's reasons for it. Proponents of JT consider it to be a strength of their view that it allows an immediately prima facie justified belief to emerge in cases where the agent typically only has their emotional experience to go on.

---

[126] See Tappolet (2014) and Deonna and Teroni (2012: 122) for what I take to be evidence that these thinkers take JT and outlaw emotion cases to be a good fit.

Therefore, it seems that proponents of JT themselves take it to be a desideratum that the epistemic role of outlaw emotions be accounted for[127]. Proponents of JT are right that JT allows outlaw emotions to provide prima facie justification to evaluative beliefs, but, in being committed to a defeater clause ED, JT risks characterizing outlaw emotion cases as ones where the emotion-based evaluative belief is never ultima facie justified.

If JT includes the defeater clause ED, emotions can only justify beliefs sans phrase if one lacks defeaters against them. But cases of outlaw emotions may be ones that we think by definition involve such defeaters, specifically rebutting defeaters, given that one has beliefs that conflict with the emotion-based belief. We, therefore, seem to have a tension between the dominant epistemic story told in philosophy of emotion literature, and the epistemic role we want to grant outlaw emotions. The objection can be summarized as follows:

> **P1:** Accounting for Outlaw Emotion cases is a desideratum of an adequate epistemology of emotions.
>
> **P2:** JT includes a defeater clause ED whereby emotions can only justify evaluative beliefs sans phrase in the absence of defeaters.
>
> **P3:** Outlaw emotion cases involve defeaters.
>
> **P4:** If P2 and P3 then beliefs based on outlaw emotions cannot be justified sans phrase.
>
> **P5:** If beliefs based on outlaw emotions cannot be justified sans phrase, then JT rules out cases of outlaw emotions.
>
> **P6:** If P1 and if JT rules out cases of outlaw emotions, then JT is not an adequate account of the epistemology of emotions.
> _____
> **C:** JT is not an adequate account of the epistemology of emotions.

This objection leaves us with a difficult choice between either rejecting JT, the dominant framework for securing emotions a normative epistemic role, or, denying outlaw emotions their radical epistemic role. I will outline why neither of these options should be accepted, and argue that the objection should, therefore, be resisted.

---

[127] They do not say so explicitly, and accounting for outlaw emotion cases has not been their concern, but in levying the success of JT at accounting for outlaw emotion cases as a strength of the view, they clearly take outlaw emotion cases to be epistemically important ones that our epistemology of emotion should make sense of.

The first option is to accept the objection's conclusion and reject JT. This can involve rejecting a component of JT, and endorsing a distinct justificatory thesis in its place, or the rejection of a justificatory epistemic role for emotions altogether. Accepting my objection's conclusion would therefore involve one of the following three options, all of which are undesirable:

1) Rejection of JT's defeater clause ED. This option would lead to emotions always providing ultimate facie justification to evaluative beliefs based on them.

2) Rejection of the immediacy component of JT. This option would mean that emotions can only justify evaluative beliefs sans phrase when the agent holds separate (justified) beliefs about the positive epistemic status of the reasons on which the emotion is based.

3) Rejection of JT altogether. This option would mean that emotions only play motivational epistemic roles.

Option 1 is too permissive. It would have all emotion-based evaluative beliefs be ultima facie justified. The justificatory force of emotions would, on such a view, admit of no defeaters. This is clearly an undesirable option as it secures emotions a stronger epistemic role than we grant other sources of epistemic justification. Option 2, on the other hand is too demanding. Option 2 would have emotion-based evaluative beliefs only be ultima facie justified when the agent holds a justified belief about the reasons one has for one's emotion. This would not only rule out outlaw emotion cases, but any case where the agent doesn't have non-emotional access to the reasons for their emotion, as well as cases where they do not have separate reason to think their emotions are properly reason-tracking. Option 3 would involve emotions playing no justificatory role at all.

The prospects of abandoning JT are, therefore, undesirable. JT still represents the best option for making sense of the epistemic role of emotions, including outlaw emotions. So, instead of accepting the objection's conclusion, we should try to resist the objection.

The alternative to accepting the objection's conclusion is to deny outlaw emotions their epistemic role. This involves proponents of JT biting the bullet and accepting that outlaw emotion cases cannot provide ultima facie justification to evaluative beliefs based on them. Maybe prima facie justification is as good as it gets in outlaw emotion cases anyway. JT allows

outlaw emotions to provide prima facie justification to evaluative beliefs based on them, which is more than a motivational epistemic role would allow them. Furthermore, the only options on the table where outlaw emotions could provide more than prima facie justification to evaluative beliefs based on them were options 1 and 2 above, which we saw should be rejected.

Biting this bullet is undesirable however. According to the objection on the table, JT both secures outlaw emotions an immediate normative role in providing prima facie justification, while at the same time blocking them from ever providing ultima facie justification, given the defeater clause and the fact that outlaw emotion cases involve defeaters. We shouldn't be content with a view that rules out outlaw emotion cases ever providing ultima facie justification to beliefs based on them. It is relevant to note that although proponents of JT do not say much about how outlaw emotions provide ultima facie justification, they do take their epistemic role not to be limited to prima facie justification (see Deonna and Teroni 2012b and Tappolet 2018c).[128] This gives us reason to believe that these thinkers would favour facing the objection head on, rather than biting the bullet.

We have seen that the options available if the objection from outlaw emotion goes through are all undesirable. We should therefore defend JT against the objection. The main way of resisting the objection is to deny that outlaw emotion cases involve genuine defeaters, i.e. to deny P3. The objection can be represented schematically, using Friedman's housewife case as an example, as follows:
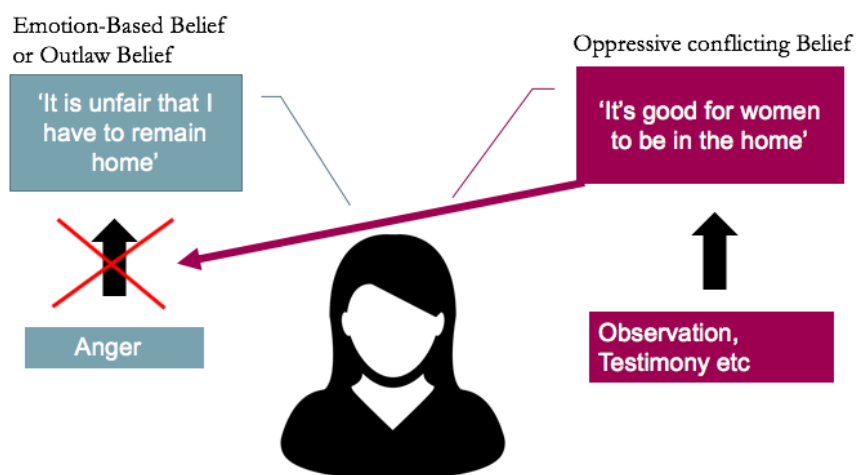

Figure 3 Schematic of objection to JT from outlaw emotions

---

[128] Deonna and Teroni (2012b:73) say that 'there appear to be judgements whose justification resists conflicting and coolly reached verdicts because the emotions that explain them respond to good reasons'.

To account for the epistemic role of outlaw emotions we need some story for why the oppressive conflicting belief does not defeat the justification of the emotion-based belief.

5.6   Responding to the Objection

I will consider two moves that can be made by proponents of JT in an attempt to deny P3, and argue that they both fail. After doing so I will argue that only a social epistemology will have the resources necessary to deny P3 and respond to the objection from outlaw emotions. This means that proponents of JT will have to import insights from social rather than perceptual epistemology to meet the objection I have raised.

Two types of move can be made to attempt a response to the objection from outlaw emotions. Both involve an attempt at denying P3. That is, both moves involve the claim that outlaw emotion cases do not involve defeaters because the would-be defeater belief is unjustified. One move is of an internalist variety, and the other of externalist variety.

5.6.1 Internalist Justification Move

An internalist response to the objection from outlaw emotions can be developed in one of two ways. Either by adding a counterfactual internalist condition such that the conflicting belief comes out as unjustified, or by making a distinction between the reasons an agent has and the reasons an agent *takes* themselves to have. I will outline each response in turn and argue that they both amount to a denial of the outlaw emotion case. As such, we will see that the externalist move is a more promising one.

The first internalist move involves endorsing the condition that the oppressive beliefs in outlaw emotion cases can only act as defeaters if they would be beliefs that the agent would endorse had they had time to reflect. In other words, only beliefs that would have survived reflection are justified and hence can act as defeaters. The thought is that Raquel, and Friedman's housewife, hold oppressive beliefs (as well as emotions and intentions) that would not survive careful reflection. This move takes the would-be-defeating belief to be one whose epistemic status the agent herself is in a position to undermine. This either amounts to a denial of outlaw emotion cases, or involves a drastic restriction of those cases

which JT would accommodate. Recall, firstly, that Raquel and Friedman's housewife both lack access to the reasons for their emotions. Reflection is likely to favour the conflicting oppressive belief given that they inhabit, and have internalized, oppressive ideology. The counterfactual condition would have to be much stronger than one that invokes reflection, it would have to say something like this: Only conflicting beliefs that do not survive reflection, personal development, research, and consciousness raising efforts can defeat emotion-based beliefs. This would be far too strong a counterfactual condition to endorse.

A second internalist move would be to hold that in outlaw emotion cases, the agent really does have most (internal) reasons in favour of the emotion, despite this not being apparent to the agent herself. The agent has more reason for the emotion, and hence the evaluative belief based on it, than they do for the conflicting oppressive belief. The agent is just mistaken about what reasons she takes herself to have, such that she forms an internally unjustified belief about the situation. The thought is that, because the oppressive belief is internally unjustified, it can't act as a defeater against the emotion, and therefore the emotion can justify the associated belief in outlaw emotion cases[129].

According to this move, agents do not necessarily have full access to their own reasons and can, therefore, be wrong about the reasons they take themselves to have. Friedman's angered housewife case would, on this view, involve the housewife having most internal reason to be angered about her situation – given that she values her own independence and has professional ambitions, for example. But these reasons, that she has, are not accessible to her. Perhaps they have been psychologically repressed by her hostile social environment, such that she makes a mistake about what reasons she takes herself to have, and continues to hold an unjustified belief that she is content with, and naturally suited for, her current lifestyle.

---

[129] This move is analogous to the one proposed by McIntyre (1990) in arguing for rational akrasia. Put simply, she argues that akrasia can be rational when, in acting against our best judgement, we are actually acting in accordance with what we have most internal reason to do. Our best judgment gets our reasons wrong because we don't have full internal access to them. Tappolet (2016) seems to adopt such an internalist move to account for emotion-based rational akrasia. It is likely she would endorse an analogous story in the epistemic case. It is at least clear that Tappolet would attempt one of the two justificatory moves, as she characterizes an outlaw emotion case as not involving defeaters: 'We might think that Padma had a reason to suspect her emotional reaction. But this thought may be resisted. The reason is that it is far from clear that Padma's Randian principles (conflicting beliefs) are justified, and it is only if they were justified that they would provide a reason to question her emotional reaction' (Tappolet 2018: 536).

An initial problem with this move is that it seems to commit a similar mistake to the one just considered. Which reasons count as ones that one possesses, but lacks access to? The sorts of concerns that count as reasons one has, but that one lacks access to, have been taken to be ones that one could have reasoned to, and over time, realized that one has (McIntyre, 1990). This move will, therefore, face the same worries as the one above. In addition to this however, this move makes the claim that in outlaw emotion cases the agent really has, at the moment of experiencing the outlaw emotion, more reason in favour of the emotion than in favour of the conflicting oppressive belief. This doesn't seem right.

Given that outlaw emotion cases involve agents that have internalized the oppressive norms of their society, it is optimistic to propose that such agents have more internal reasons for their emotion and its associated judgement than for the conflicting oppressive judgement. Of course, the agent must have *some* internal reason for the emotion – the frustration of a related desire, for example – otherwise the occurrence of the emotion in that individual is entirely mysterious. But it is implausible to use this trivial fact to infer that the agent has *more* reason for her emotion than for the conflicting judgement, at least in so far as this is meant as a general account of what is going on in outlaw emotion cases. Outlaw emotions are exactly emotions that go against the social framework the agent has internalized. The conflicting oppressive belief is the manifestation of one component of the internalized framework, and it is therefore likely to better cohere with the other beliefs, emotions, and intentions that stand in line with that framework. Such cases are best described as ones where the agent indeed has most internal reason to hold the oppressive conflicting belief.

Both internalist moves, therefore, either involve a denial of outlaw emotion cases, or a drastic restriction of those outlaw emotions that will be able to play justificatory roles. Given that outlaw emotion cases are ones that involve the agent feeling her emotion as 'outlaw', these will be cases where the agent's oppressive beliefs are internally justified. It, therefore, seems unlikely that an internalist story will be able to deliver on the denial of P3. It seems we must instead appeal to an externalist story to vindicate the epistemic role of outlaw emotions.

5.6.2 Externalist Justification Move

Friedman's housewife believes that her place is in the home and holds a number of beliefs and intentions in line with this, as well as dispositions to feel emotions such as pride when

she succeeds at fulfilling this role. What justifies her outlaw emotion of anger must be something about how the emotion tracks the external fact that her situation is an unjust one. An externalist attempt at denying P3 therefore seems more promising. Indeed, it has been argued that externalism is the preferred epistemological framework for dealing with conditions of oppression, as only external facts can vindicate the beliefs of those who have internalized the oppressive ideology surrounding them (Srinivasan in progress).

The external justification move takes the conflicting oppressive judgement not to be a defeater because it is externally unjustified. The judgement is unjustified because it is not reliably responsive to external reasons. The thought is that living under conditions of oppression puts one systematically out of touch with truths, and compromises one's belief forming capacities such that one is not reliably connected to facts (Srinivasan in progress). Considerations of the sort 'a women's place is in the home' are false, and the product of systematically misleading conditions of oppression. The processes that lead to such beliefs are therefore unreliable, making such beliefs unjustified. The oppressive conflicting beliefs, therefore, cannot, according to this move, act as defeaters against the justificatory force of outlaw emotions. P3 is, in this way, denied and the objection to JT from outlaw emotions is resisted.

Although this is intuitively appealing, the externalist move is actually harder to establish than one might think. First of all, what seems to be driving the force of this move is a commitment to the falsity of the oppressive beliefs, rather than any problem with the mechanism by which they are arrived at. In arriving at the oppressive beliefs, is the agent not treating the evidence available to her adequately, in a fashion that we would think reliably generates true beliefs had the beliefs been true? Is the agent not reasoning correctly from the testimony of otherwise reliable agents, and from her otherwise reliable empirical observation? It seems that the issue is not the reliability of her belief forming processes per se, but rather, as I will suggest below, their relation to a social structure in which the agent is embedded. What drives the plausibility of the externalist move then, is the fact that the beliefs generated tend to be false, this is what gives us reason to think the belief-forming processes involved must be unreliable.

This move seems to get something right, after all it is false ideology that outlaw emotions are thought to provide evidence against, and people living under conditions of oppression get

things systematically wrong about the world by believing all sorts of insidious falsities. But isn't there a sense in which the statement 'a women's place is in the home' gets things right about the society in question? Similarly, isn't there something true about statements like 'women are submissive' or 'women shave' when uttered within the context of present-day patriarchal societies? Indeed, isn't it exactly one of the features of such oppressive societies that stereotypes become self-fulling and self-perpetuating? Such considerations suggest that beliefs with ideological content could be both justified and true. If this is the case, then it would seem that even the externalist attempt at denying P3 fails, since the conflicting oppressive belief may still be best construed as a potential defeater to the emotion-based belief. This means that the conclusion of the objection could stand, and we may be forced to either abandon JT, or to deny that outlaw emotions can justify their associated evaluative beliefs.

Let me say a bit more on the plausibility of self-fulling oppressive social truths. The prevalence of self-fulling oppressive facts is widely acknowledged, and argued for, in feminist philosophy (Frye 1983; Mackinnon 1987; Langton 2009; Haslanger 2011). McKinnon (1987), for example, writes that oppressive beliefs 'become proven, in part because the world actually arranges itself to affirm what the powerful want to see' (164), similarly, for Langton (2009):

> When oppression is systematic enough, there is nothing accidental about a correlation between beliefs that women are servile, and women's servility; and the connection between beliefs and the truth of those beliefs can be as reliable as one could wish. (303)

Oppressive beliefs can, therefore, be true and justified. This doesn't mean that these beliefs are not problematic, however, just that their problem need not be epistemic. Langton takes self-fulling oppressive beliefs to not only be instrumentally harmful, but to constitute harms against those oppressed. That such oppressive states of affair have become true is a harm. In a similar vein, Haslanger (2011) argues that generics like 'crop tops are cute', 'women are submissive' and 'cows are food' can be true, and beliefs based on them justified. Haslanger (2011) writes that:

> Under conditions of male dominance, women are, in fact, more submissive than men. This is a true generalization and those who live under male dominance are justified in believing it. (182)

Haslanger locates the problem with such generics not in their semantic truth value, as they adequately capture the way the social world happens to be arranged, but in their pragmatic

dimension: specifically, in that these generics tend to imply that it is a necessary part of the nature of women, crop tops and cows to be submissive, cute and food, respectively. Women are not by their nature submissive, nor are cows 'meant for' human consumption. Essentialist claims of this sort naturally generate normative claims, such as that women 'should be submissive', for example. The descriptive generics should, therefore, be resisted not for lacking epistemic credentials but because they tend to pragmatically imply, as opposed to logically imply, essentialist and normative beliefs which are both false and morally problematic. Our descriptive oppressive generics can be true and justified according to Haslanger, despite us having reason to resist them given their pragmatic implicature of false normative and essentialist views.

In outlaw emotion cases then, beliefs such as 'women are home keepers', can be based on social facts present in the oppressive society in question. These social facts presumably act as reasons that can justify the generic beliefs. Outlaw emotion cases might therefore involve true, albeit problematic, externally justified oppressive beliefs, that are formed by a reliable process.

At this point the proponent of the externalist move to deny P3 might claim that I have only managed to establish that *descriptive* oppressive beliefs can be justified, and remind me that those oppressive beliefs which are would-be-defeaters in outlaw emotion cases are typically *normative*. Recall that the oppressive belief in my schematic of the Friedman's housewife case was that 'it is *good* for women to be in the home', while Raquel's sexual harassment case involved the oppressive belief that 'I *should* be flattered by the man's actions'. These are normative beliefs. The thought is that, in so far as some descriptive oppressive generics might have positive epistemic value, their normative counterparts do not, and as the oppressive beliefs in outlaw emotion cases are typically normative, their epistemic status cannot be vindicated. The proponent of the externalist move to deny P3 can therefore discard the insights from Haslanger as orthogonal to their move.

This would be too quick however. There is room for a defense of the epistemic standing of even the normative oppressive beliefs. Even if we take the normative oppressive beliefs to be false, they might still be justified, which would block the externalist move against P3. Indeed, false oppressive beliefs are presumably formed by the same mechanisms that led to the justified and true oppressive descriptive beliefs. If descriptive oppressive beliefs are

justified when they are reliably connected to features of social reality, then we have reason to think that normative oppressive beliefs are formed by similarly reliable mechanisms and, hence, can be justified. There will surely be evidence in favour of thinking not only that that women are submissive but that women should be, and are by nature, submissive, under conditions of patriarchy. Langton (2009) introduces a distinction that is helpful on this matter:

> In conditions of oppression, expectations of how a subordinate will behave can be self-verifying or self-fulfilling… if a belief is self-verifying, it provides evidence for itself (given a certain context). If it is self-fulfilling, it makes itself true (again, given a certain context). Beliefs that are self-verifying may be false, but justified: the evidence may be misleading, but being evidence, it supports the belief, even if the belief is false. (300)

The thought is that oppressive beliefs can be self-fulling, when they make the world conform to them, but they can also be self-verifying, when they make the world provide evidence for them. False oppressive beliefs with normative content may, therefore, often be justified, in so far as they reliably correlate to social evidence. This would mean that Raquel, as well as Friedman's housewife, have justified, albeit false, beliefs that 'it is good for women to be in the home', and that 'I should be flattered by the man's actions', respectively. This means that the externalist move is still challenged, as the conflicting beliefs in outlaw emotion cases seem to come out as justified, no matter whether they are true[130].

Given the above, attempts to deny P3 that hinge on outlaw emotion cases not involving justified oppressive beliefs, face difficulties in accounting for epistemic life under oppressive conditions. I will sketch an alternative way of resisting P3 that is faithful to the considerations

---

[130] For the defender of the externalist move to push back against this, they would have to claim that defeaters in outlaw emotion cases must not only be justified but true. Unfortunately, I think there is likely even room to argue that oppressive normative beliefs might sometimes turn out to be in a sense true as well as justified. It might be a fact within the patriarchal society of Friedman's housewife that it is good for women to stay in the home. The social practices and beliefs of that society make it the case that it is good, and better than other options, for women to stay in the home. Similarly, it might be true that Raquel's experience is reason for flattery, within her social context. There seem to be two senses of truth at play: moral truths as relative to a moral community, and truth regarding how things ought to ideally be, or indeed are 'deep down'. Many find it important to take seriously both types of truth, without discarding truth relative to moral communities as epistemically illegitimate (Calhoun 2004; Haslanger 2012). In any case, it is hard to deny that it is a (descriptive) truth that the normative beliefs 'it is good for women to stay in the home' and 'Raquel should to be flattered' are true of their respective moral communities. There are therefore likely to be true descriptive beliefs with normative content. There is therefore room to undermine a move at denying P3 that restricts defeater status to those beliefs that are both justified and true - A move which many would find too demanding in its own right to begin with.

above. I will suggest that adopting an overtly *social* epistemology is the most promising way of saving JT's main insight, while allowing outlaw emotions their epistemic role.

5.7   Alternative sketched

When we take ideology seriously then, I have suggested, it isn't easy to establish that the would-be defeater in outlaw emotion cases is unjustified. My sketch of an alternative way of denying P3 will not hinge on denying the would-be-defeater its epistemic standing. To begin with, I will assume a form of externalism of the sort arrived at above. An externalism that takes ideological statements to often be true and justified is one that properly maps a troubling feature of life and knowledge under conditions of oppression. In line with this, and externalism in general, I will assume some form of epistemic reliabilism, where a belief is justified if it is produced by a process that reliably generates true beliefs. According to my view therefore, emotion-based beliefs, as well as beliefs arrived at non-emotionally, via authoritative testimony or deliberation, for example, are reliably formed, as these processes are reliable at generate true beliefs. This will, therefore, be true even of many self-verifying or self-fulling ideological beliefs, such as 'it's good for women to stay in the home', when these beliefs occur within the relevant oppressive contexts, and are arrived at through processes that reliably respond to evidence in their favour.

With these preliminaries in place, let's now turn to outlaw emotion cases. In these cases, we seem to have two conflicting beliefs that are prima facie justified, one based on the outlaw emotion and the other informed by the agent's social context. The belief that 'my place is in the home' is not only internally justified by a number of other ideological beliefs with which it coheres, but may also be externally justified by statistically relevant social patterns picked up by reliable processes such as observation and reliable testimony. The belief that 'my situation is unjust', which is based on the emotion, is also prima facie justified, given our ongoing assumption that emotions are reliably correlated to the generation of true evaluative beliefs. The problem is why the conflicting oppressive belief doesn't act as a defeater against the emotion-based belief.

5.7.1 A Social Epistemology Framework

My alternative story for why this is so involves combining the insights of reliabilism and standpoint epistemology such that the objection from outlaw emotions can be met[131]. A thorough consideration of standpoint epistemology lies beyond the domain of this chapter[132]. For our purposes we will take the main insight from standpoint epistemology to be the claim that oppressed individuals are in a position of epistemic privilege regarding their own oppression[133]. The claim is not that being a member of an oppressed group is necessary or sufficient to gaining knowledge about the relevant domain of oppression, but rather than oppressed groups are more likely to gain knowledge about their oppression given their particular social standing[134]. This is because social positions make a difference to what evidence agents have, as well as their dispositions to recognize evidence as evidence (Toole, forthcoming).

What seems natural to suggest then, is that one's social standing makes one a more reliable truth tracker relative to a certain domain. So, women will more reliably track the reality of society's gender structures, while people of colour will more reliably track truths regarding racial relations. Reliability here is comparative, the oppressed are more reliable, as compared to those who are not oppressed, at tracking truths relative to a relevant domain. Kitcher (1994) has made a similar point and says that 'reliabilists should insist that some standpoints are better or worse than others with respect to certain types of propositions' (125).

We therefore have a reliability claim about social standpoints:

---

[131] For the compatibility of standpoint epistemology and process reliabilism see Michaelian (2008)

[132] See Hartsock (1983) and Harding (2004)

[133] Here I am taking the fact that oppressed individuals are oppressed to grant them some epistemic privilege regarding the nature of this oppression. This is the case despite their social position typically falling short of a *feminist* standpoint, for example, as many take such standpoints not to be inhabited but achieved. Some reserve epistemic privilege only for standpoints that are achieved, rather than the social positioning of those agents prior to the standpoint being achieved (Toole forthcoming). I do not. Consciousness raising for example can be instrumental at getting an oppressed woman to achieve a feminist standpoint. The woman's social situation prior to achieving the standpoint however, made her more disposed to achieve the standpoint. Of course, once she occupies a feminist standpoint she will have even more epistemic privilege relative to gender oppression than before, but I take her to have inhabited a position of (dispositional if you like) epistemic privilege all along. Throughout, my use of the term 'standpoint' is meant to refer to this weaker position of epistemic privilege, as outlaw emotion cases are critical to achieving the feminist standpoint and therefore do not arise, in the first instance, from within it.

[134] This is a distinction between modernist or Materialist standpoint epistemologies such as Hartsock's (1983) and the post-modern or Social standpoint epistemologies such as Harding's (2004). The former are charged with holding a necessary connection between epistemic privilege, labour and group membership, while the former allows a plurality of experiences within the same standpoint and takes the connection between epistemic-privilege and group membership to be looser (Michaelian 2008; Toole forthcoming).

Standpoint Reliability (SR): Oppressed individuals are more reliable at generating true beliefs relative to particular domain(s) of oppression.

What specific features of one's social position grant one epistemic privilege? Marxist standpoint epistemologists take women's contact with the material conditions of reproductive labour to underlie their epistemic privilege regarding gender oppression. Most contemporary standpoint epistemologies have, however, moved away from strict materialist readings and ground the epistemic privilege of groups in a wider set of relations they bear to surrounding social conditions and how these structure their experience (Fricker 1999; Dotson 2012; Toole forthcoming). There has, therefore, been a shift from focusing on the relation groups bear to forms of labour, to focusing on the social experiences of marginalized groups.

Outlaw emotions are one relevant experience of marginalized individuals. We saw a similar commitment in Jaggar (1989) who wrote that 'people who experience conventionally unacceptable, or what I call 'outlaw' emotions often are subordinated individuals who pay a disproportionately high price for maintaining the status quo' (166). Outlaw emotions are sensitive to evidence that oppressed agents are in a privileged position to access. This is partly because one's dispositions to experience particular emotions depends on one's social standing. Standpoints will, therefore, dispose agents to be sensitive to particular affective reasons. One type of belief-generating process by which oppressed individuals are more reliable at generating true beliefs (SR) then is emotional. As Kitcher says (1994) 'the claim that a particular standpoint is preferable to others can thus be recast in terms of the relative reliability of the processes that different standpoints make available' (124). This gives us a second claim:

Standpoint Emotions (SE): Occupying a particular standpoint makes one more sensitive to particular affective reasons.

Outlaw emotions are, therefore, one of the belief-generating processes made more easily available by occupying a particular standpoint, because occupying a standpoint will dispose one to greater sensitivity to certain affective reasons (SE). Emotions are, however, not the only belief-generating processes that standpoints make available. A doxastic version of SE is

likely true as well. That is, occupying a particular standpoint makes one more sensitive to particular epistemic reasons. It is plausible that members of oppressed groups, therefore, are also more reliable at generating true beliefs about their oppression, through other means, such as reasoning for example. Do we have reason to think that emotions constitute a particularly reliable belief-forming process, within agents that occupy privileged standpoints? Indeed, we might.

We saw (in sec 5.2.2. as well as in 4.4) that affective reasons are strict and demonstrate sufficiency. This means that they have justificatory force that cannot be cancelled out by the presence of other affective reasons, nor by just any non-evaluative epistemic reason. The features that characterize the irreducible reason-relation distinctive of emotions might allow them to more reliably track truths under conditions of oppression.

Conditions of oppression are fertile grounds for the double consciousness that outlaw emotion cases are often involved in. In such cases, one has justified beliefs that conflict with each other. In such cases, one typically has a lot of evidence in favour of oppressive beliefs which count against, and often cancel out, epistemic reasons in favour of outlaw or progressive beliefs. Reasons for emotions, however, do not cancel each other out, nor are they undermined by insufficient non-evaluative epistemic reason against them. This means that in situations where there are (evaluative and non-evaluative epistemic) reasons 'against' a certain emotion-based belief, the emotion will be more reliable at tracking certain reasons than beliefs. This is because emotions can better track reasons in the face of disconfirming evidence, than beliefs. This makes emotions a particularly reliable belief-generating processes under conditions of oppression. This gives us the following claim:

> **Emotional Reliability (ER):** Emotions are more reliable than non-emotional processes at generating true beliefs under conditions of oppression.

We have seen that one intuitive way of making sense of SR is to cash out the reliability of standpoints in terms of the reliability of the belief-generating-processes it makes available. Emotions constitute one such belief generating process. This gave us SE, the claim that standpoints make us more prone to track particular reasons through emotional experiences. ER states something further, which is that, within the processes made available by a privileged standpoint, emotions constitute a particularly reliable belief-generating process

given the distinctive reason-relation characteristic of emotions. This is because their sui generis normative rationality makes emotions particularly well-suited to tracking reasons in the face of disconfirming evidence.

5.7.2 Responding to the Objection from Outlaw Emotions

We are now in a position to turn to how the outlined framework is better equipped to respond to the objection from outlaw emotions than the attempts we considered above. What does a denial of P3 look like on my account?

The outlaw emotion cases that we have been concerned with involve agents that, according to SR, occupy privileged standpoints relative to domains of oppression. Friedman's housewife and Raquel are both women living under conditions of extreme gender oppression. According to SE, they are therefore agents whose social position will have played a role in disposing them to experience the outlaw emotion. Furthermore, given ER, their outlaw emotion is more reliable at generating true beliefs. The conflicting oppressive beliefs in the cases of Raquel and Friedman's house-wife are, however, also generated by processes available to agents occupying privileged standpoint, after all Raquel and Friedman hold those oppressive beliefs despite being, following SR, in a position of epistemic privilege.

Why doesn't the oppressive belief defeat the justification of the belief based on outlaw emotions in such cases? I take it that ER alone is insufficient to ground a rejection of P3. A denial of P3 based solely on ER would say that because emotions are more reliable under conditions of oppression, the conflicting oppressive belief cannot act as a defeater against it. This is an undesirable outcome however. This would result in an endorsement that all beliefs based on emotions occurring under conditions of oppression, remain undefeated by conflicting beliefs that are not based on emotions. This is not a good outcome seeing as oppressive conditions shape many of our emotional experiences and ER does not differentiate between oppressive emotions and outlaw emotions. This would mean that one's fear of ethnic minorities could justify the belief that these ethnic minorities pose a threat, even in cases where one holds an explicit belief to the contrary. If only ER is appealed to, this would be the case even if this fear is experienced by a member of the very ethnic minority that is the object of fear. Furthermore, appealing only to ER would not be able to account

for cases in which the oppressive belief is itself based on an oppressive emotion. If Raquel feels proud and angry at the actions of the man who squeezed her, appealing only to ER will not be able to differentiate between the epistemic standing of these emotions or the beliefs based on them.

I propose that we need to make reference to the content of the beliefs in question, rather than merely the processes that generate them, in a denial of P3. Specifically, we need to make reference to which standpoints beliefs with particular contents most typically arise in. Following SR, women are more reliable knowers than men regarding gender oppression, for example. So, women are more likely than men to form the belief Friedman's discontent housewife forms (i.e. that the situation is unjust). I propose that P3 can be denied then, by appealing to the fact that the conflicting oppressive belief is one that would feature disproportionately in members of groups that lack epistemic privilege, as compared to those that have epistemic privilege relative to that domain. The thought is that the conflicting beliefs in outlaw emotion cases are not defeaters because they are less reliably connected to truth than the emotion-based beliefs. Beliefs based on outlaw emotions are more reliable because they have *content* that is more likely to feature in the beliefs of those who occupy a position of epistemic privilege relative to that domain.

This is different to the way in which the externalist move considered in section 5.6 attempted to deny P3. The externalist move above denied oppressive beliefs justification. On my account the oppressive conflicting belief is typically justified. The conflicting belief is, however, less reliable than the emotion based one and therefore cannot defeat it. But this is not due to the belief having been formed by a defective *process*. This is because the belief has *content* that is more likely to be shared by agents that are less reliably connected to truth, because they do not occupy a position of epistemic privilege relative to the domain of oppression.

This story is immune to concerns that racist emotion-based beliefs will come out ultima facie justified in the face of a conflicting belief. This is because the conflicting belief will defeat the emotion-based belief. The conflicting belief can do so because it is more reliable than the emotion-based belief, given that it is a belief more likely to feature in the beliefs of those groups that are epistemically privileged regarding racial oppression, compared to those of groups that are not.

My alternative story can also capture how the outlaw emotions of individuals that are not members of an epistemically privileged group can form ultima facie justified beliefs based on them. Take Huckleberry Finn, for example. If he forms the belief that Jim deserves respect (and hence doesn't deserve to be turned in to his owners), based on his feeling of respect for Jim, Huck forms a prima facie justified belief. Now, Huck also holds explicit beliefs that Jim does not deserve the respect his outlaw emotion urges him to grant Jim, after all Huck knows Jim is the property of a slave-owner, and believes he should turn Jim in. Why doesn't Huck's oppressive beliefs defeat the justification of his emotion-based belief? Huck is white and, therefore, does not occupy a position of epistemic privilege with respect to the relevant domain of oppression. But his emotion-based belief has content that is more likely to be shared across non-white individuals, i.e. those who occupy a position of epistemic privilege relative to the domain of racial oppression. His emotion-based belief is therefore more likely to be true because those in positions of epistemic privilege are more reliably correlated to truth. His oppressive belief, therefore, does not defeat his emotion-based belief.

5.7.3 Two Worries

A worry for my alternative response to the objection from outlaw emotions, that might be called 'Male Feminists', goes as follows: What if enough men have become feminists such that there is no difference between the beliefs of the group that occupies a position of epistemic privilege and the group that does not? A second worry that could be called 'Sexist Women' is as follows: What if those occupying the position of epistemic privilege just fail to have feminist beliefs and instead have sexist ones?

A first point in response to 'Male Feminists' is that it essentially amounts to a denial of the outlaw emotion cases. If the situation described in 'Male Feminists' obtained then one's outlaw emotion would be unlikely to be felt as outlaw, after all, many people within one's community would hold similar progressive beliefs, and the society in question would just be radically different to the one in which the outlaw emotions we have been concerned with arise. That being said, we could grant, for the sake of the argument, that there are outlaw emotion cases occurring in societies where exactly the same number of women and men hold a belief that matched the content of one's outlaw-emotion-based belief. How then do

we establish that one's outlaw emotion more reliably tracks truth if we cannot appeal to some beliefs being more prevalent in certain standpoints than others? One move here would be a genealogical one, that is, we could refer to the standpoint that generated the evidence that allowed the dominant group to become aware of certain truths. The epistemic privilege of the oppressed group would still be salient on this sort of genealogical story, and one could make reference to it in denying P3. P3 would be denied because the emotion-based belief is more reliable that the would-be-defeater – it is more reliable because (even though it is shared across other standpoints) it originated from an epistemically privileged standpoint.

We could also invoke the fact that according to my view (fn. 133), women stand in a privileged position regarding their *capacity* to occupy the feminist standpoint. So, a response based on capacity, rather than actuality, could be developed. Such a response would have the emotion-based belief be more reliably related to truth than the oppressive belief, because it is one endorsed in the feminist standpoint, while the oppressive belief is not. The feminist standpoint is one which women have a greater capacity to occupy than men, whether they indeed occupy it or not. That is, women are in a privileged position with respect to their capacity to occupy the feminist standpoint.

The situation described in 'Sexist Women' is more worrying, partly because it is more realistic. A similar capacity-based response could, however, be developed to dissolve the worry. That is, women, even if sexist, are in a privileged position regarding their capacity to achieve a feminist standpoint. And the beliefs characteristic of this standpoint could be made reference to in an effort to undermine P3.

If the 'Male Feminists' and the 'Sexist Women' cases are thought experiments so abstract as to deny that certain groups have privileged capacity to occupy certain standpoint, then they deny something I have taken to be a fact about epistemic life under oppression, namely SR. I'm happy to concede that the alternative story I have provided for denying P3 doesn't work in cases where SR comes out as false, as these will be far removed from social reality as we know it, and far removed from the situations in which outlaw emotions occur.

A last worry to mention, that arises in my response to the two worries just outlined, is that my story involves having to adjudicate between epistemically privileged and unprivileged beliefs. That is, the content of beliefs, as opposed to features of their generation, will have

to be evaluated. This is indeed a demanding feature of my account, that might be a downside to it, but it seems like a small price to pay compared to previous attempts at denying P3, which we saw either risked ruling out outlaw emotion cases, or significantly misconstrued epistemic life under oppression.

5.7.4 Alternatives to my alternative

Why not just drop the social epistemology framework and deny P3 by appeal to the outlaw emotion's relation to truth? We could presumably attempt a denial of P3 that hinges on outlaw emotions being more reliably related to truth, in a manner that does not make reference to social standpoints. What would such an account look like?

We saw in section 5.7.2 that invoking only ER would lead to the undesirable consequence of any emotion occurring under conditions of oppression remaining undefeated in the face of conflicting beliefs. To bolster this view and avoid this consequence, one would have to invoke SE as well as ER. This alternative would presumably hold that only those outlaw emotions that one has *in virtue of* occupying the standpoint one occupies, should remain undefeated by conflicting beliefs. But in doing so this alternative already imports features of a social epistemology. It does so while proving less explanatorily powerful than my alternative, for it excludes cases such as Huckleberry Finn's. As Huck does not have the relevant outlaw emotion in virtue of occupying his standpoint, he has it *despite* his unprivileged standpoint.

Why not just say that conflicting beliefs cannot defeat beliefs based on outlaw emotions when these emotions are justified? This would mean endorsing a view whereby all *justified* emotions provide ultima facie justification to the beliefs based on them. There are a few problems with this. Such a move would mean that any case where there are affective reasons for one's emotion, and hence the emotion is justified, an associated evaluative belief based on it is *ultima facie* justified. This means that beliefs based on justified emotions do not admit of defeaters. This seems too strong. We might think there are cases where emotions are justified but evaluative beliefs based on them would not be ultima facie justified. If your friend offends you and you become angry, your anger is justified and your belief that they have offended you is prima facie justified according to JT. But let's say your friend offended

you because they are themselves going through a hard time. There is some excuse that defeats the ultima facie justification of a belief that your friend has intentionally offended you. This excuse need not defeat the fact that you have reason to feel offended. In line with this, we saw that there can be insufficient reason for evaluative beliefs in scenarios where there is sufficient reason for an emotion. Recall your suspicion of Maria. It might be justified, despite you not having sufficient *epistemic* reason for the belief that she is untrustworthy. You have prima facie reason for a belief that Maria is untrustworthy, given your feeling of suspicion, but it would be undesirable if this belief came out as immediately ultima facie justified, and did not admit any defeaters.

There is reason to resist a view whereby all justified emotions provide ultima facie justification to the beliefs based on them, that does not rely on previous arguments I have made. This is that such a view will still need to provide a story for why the conflicting oppressive belief doesn't act as a defeater against the justification of the outlaw emotion itself. That is to say, the reasons in support of the emotion will have to remain undefeated such as to be sufficient to provide ultima facie justification[135].

Although there is still much to work out, the alternative I have sketched grants that both the outlaw emotion-based judgement as well as the conflicting oppressive belief can be justified. We have seen that making sense of the conflicting belief as justified (internally and externally) is important if we are to be faithful to life and knowledge under oppression. The thought, therefore, is that beliefs based on outlaw emotions can be justified because of their relation to a privileged, and hence more reliable, standpoint than the would-be defeaters with which they conflict. My sketched alternative involves a denial of P3, as the conflicting belief doesn't defeat the emotion's justificatory role. My alternative, unlike those discussed above, involves

---

[135] Indeed, Deonna and Teroni (2012b) seem to hold that evaluative judgements based on justified emotions are ultima facie justified. It is not entirely clear whether they think the evaluative judgements admit of defeaters that would not apply to emotions in certain cases, but it *is* clear that they take outlaw emotion cases to be ones where a story must be told for why the conflicting beliefs do not defeat the justification of the emotion-based evaluative belief. They propose very briefly the condition that: 'the subject's judgement is justified if, absent the factors clouding his coolly reached conclusion, he would realize that the reasons his emotion responds to defeat his reasons to judge otherwise' (2012b:fn.42). This can be read as amounting to the following denial of P3 in outlaw emotion cases: oppressive conflicting beliefs do not act as defeaters because if oppressive conditions had not been the case one would not have formed the belief. But this would amount to saying that conflicting beliefs cannot defeat the justification of beliefs, if the conflicting beliefs would have been unjustified had there been no evidence in favour of holding the conflicting belief. This is trivially true and doesn't do much as a response to the objection from outlaw emotions. Another reading of Deonna and Teroni's condition would be, that absent oppressive conditions, the agent would be able to judge that the reasons for one's emotion defeat the reasons for one's oppressive belief. But absent oppressive conditions outlaw emotion cases do not arise in the first place.

a denial of P3 that proves plausible because it does not stand in tension with important features of emotional life under oppression.

5.8   Conclusion

Our fourth desideratum was to provide an account of anger's epistemic rationality. I began by defending a justificatory epistemic role for emotions, such that in being felt, anger can provide immediate prima facie justification to evaluative beliefs. This matches the epistemic role anger seems to, at least sometimes, play in our personal lives. It also matches the epistemic role that feminist philosophers have taken anger to play. I defended this justificatory epistemic thesis (JT) against objections made against it in the philosophy of emotion literature. In doing so, I provided a novel, and stronger, response to the Superfluity Objection, one that takes steps towards securing emotions a distinctive normative role as opposed to a dispensable one. One type of case where emotions play distinctive epistemic roles is in outlaw emotion cases. JT seemed initially well equipped to account for outlaw emotion cases, but I argued that JT actually risked ruling such cases out given it's neglected defeater clause ED.

I developed a novel, and potentially disrupting, objection to JT in light of this: the objection from outlaw emotions. I argued that the obvious moves that might be made in response to the objection prove unsuccessful. In doing so, an externalist social epistemology emerged as the preferred framework for making sense of outlaw emotion cases. My alternative proposal for denying P3 involved marrying features of realiablism to features of standpoint epistemology.

Although many specifics are left to be developed - how does one identify beliefs generated by a privileged standpoint, for example? - I have provided the foundations for an alternative framework that seems to be the most promising way of responding to the objection from outlaw emotions. I have, therefore, taken significant steps in the direction of securing a strong and distinctive epistemic role for anger under conditions of oppression.

By providing a systematic account for the positive epistemic role of emotions, I have supplemented the arguments in favour of anger's epistemic value that remained programmatic in the work of those who praised anger. In so doing I have taken significant

steps in meeting my final desideratum. Given my account, anger has emerged as a state capable of playing not only strong, but crucial, epistemic roles. This suggests that anger is epistemically rational. Those who wish to dispute this will have to supply either challenges to the argument I have made, or, reasons for thinking anger has negative biasing effects, in causing agents to form false beliefs that favour their interests for example, that outweigh the positive role I have argued for. In section 5.2.4 we saw empirical evidence that suggests that the biasing effects of emotions are likely to have been exaggerated. Whether disconfirming evidence can be provided on this or not, I have set the bar for undermining anger's epistemic rationality significantly higher. This is because I have argued that anger not only plays positive, but crucial epistemic roles.

6

The Rationality of Anger: Concluding Remarks

*'What I really mean to say is that we think and talk about political anger
in the way we do because it serves those whom anger
most stands to threaten, and that this is no mistake at all.'*
-Srinivasan, in 'The Aptness of Anger'

My contribution to the debate on anger has been, in the first instance, methodological. I argued that a foundational approach to the debate was called for. An approach that makes the emotion of anger central was in order for three related reasons. Firstly, because we should have a deep, rather than cursory, understanding of the phenomenon in dispute. Secondly, because a foundational approach allows points of substantive disagreement within the literature on anger to be more clearly made out, and lastly, because such an approach promises to help discern which side of the debate should be favoured.

My substantive contribution to the debate has been the provision of an account of anger's nature and rationality. In doing so, I provided novel arguments in favour of the political and epistemic value advocated of anger by feminist philosophers. I argued that the points over which disagreement arises in the debate on anger all relate to some form of the emotion's ir(rationality). Disagreement arose over whether anger was a benefit or a hindrance to 'rational thinking', which I argued relates to concerns over the emotion's normative, and epistemic, rationality. Further disagreement arose over whether anger was beneficial to securing the agent's practical aims, which I took to involve disagreement over anger's instrumental rationality. Finally, there was disagreement over anger's nature, which impacts not only what types of rationality anger is capable of, but whether disagreements in the debate have the same target phenomenon in mind to begin with. This generated four desiderata for moving the debate forward:

1. The provision of an account of anger's nature
2. The provision of an account of anger's instrumental rationality
3. The provision of an account of anger's normative rationality
4. The provision of an account of anger's epistemic rationality

These are desiderata that in jointly being met provide a robust, albeit non-exhaustive, account of anger's rationality. In pursuing them, the traditional conception of anger's nature as involving a constitutive desire for retribution was challenged, and attempts at dissolving the debate by appeal to distinct referents were set aside. Anger emerged as a heterogeneously constructed kind, characterized by instrumental rationality, sui generis normative rationality, and epistemic rationality.

The traditional construal of anger as constitutively tied to a desire for retribution emerged as not only overly simplistic, but potentially biased. We saw that the 'basic' construal of anger from the affective sciences, that could be levied in support of a traditional conception of anger, did not survive scrutiny. Human anger is not its affect program. By this I do not mean that anger is not *just* its affect program, but that in having been subjected to various stages of cooperative adaptive pressures, and having involved exogenetic inheritance associated with these adaptive conditions, human anger does not involve affect programs in the common sense of the term. I proposed that the anger affect program is best seen as the evolutionary *ancestor* of anger phenomena shared across the human population, as well as some animals.

The strict link between anger and particular actions was weakened, and its tie to any unique function questioned. I argued that, even if being constitutively tied to retribution admitted a massive plurality of angry behaviours, we have empirical, as well as philosophical, reasons to doubt a constitutive link between anger and the desire for retribution. I provided a novel, experimentally informed, head-on defense of anger against the *inefficacy claim*. This was the claim that anger is typically ineffective or counterproductive at securing the aims of angry agents. Defenses against this charge have been scare in the philosophical literature because it depends partly on empirical claims. Empirical work, as well as the dynamics of apology, highlighted the desire for recognition as central to anger. It emerged that, as tied to a desire for recognition, anger is far less susceptible to charges of inefficacy or instrumental irrationality. Indeed, I argued that in often aiming for recognition, anger is prima facie instrumentally rational in personal and political life. I focused on anger's role in fighting against social injustice. I did so because this is the role that those who praise anger have most forcefully advocated for. Recall that I restricted my examination of anger's instrumental rationality to cases of *constructive* beneficial effects. That anger can be efficient at securing, at the very least long-term, social change through destructive and punitive actions is uncontroversial, although the ethics of such efficacy is, of course, complicated. By arguing for anger's constructive efficacy, I side-step ethical worries and provide a challenge to views that take anger's efficacy to come at the expense of ethical complications.

The empirical work I relied upon allowed us to see anger that is tied to retribution, as one way in which anger can manifest, rather than as anger's natural, or even prototypical, core. Indeed, I suggested that anger's tie to retribution might be the exception to the norm. There is empirical evidence in support of the view that anger only involves desires for retribution under circumstances where agents have 'nothing to lose'. This would suggest a near complete inversion to Nussbaum's (2015, 2016) picture, where non-retributive anger is exceptionally rare. My challenge to the traditional conception of anger's nature does not depend on this strong claim, however. I need only endorse the more modest thesis that anger typically involves one of two distinct desires: for retribution, or recognition, to have challenged the traditional, purely retributivist, view.

At present, the account offered of anger's nature is explanatorily superior to the traditional view. It can explain the prevalence of the traditional retributive view of anger, while not reducing anger to it. The account also fits well with a range of evolutionary, neuroscientific

and psychological evidence, which the traditional view seems ill equipped to account for. I suggested that the retributive view of anger may have gained prevalence precisely because conditions that favour retributive manifestations of anger were typical of societies with deeply engrained social hierarchies. The thought was that two key moderators, that affect whether anger is associated with a desire for retribution over recognition, are strongly entrenched in these societies. My account lends support to the view that our traditional conception of anger is biased by those that benefit from such a conception. It is in the interest of those in power to characterize anger as being by nature vindictive and irrational. It would be in the interest of those in power to dismiss the anger of those in inferior positions as normatively irrational, that is, as inappropriate. Perceptions of appropriateness was one of the key moderators I highlighted. By having a vested interest in the dismissal of the anger of subordinate groups, then, those in power purposefully perpetuate this moderator. This, in turn, makes retributive manifestations of anger more likely, which 'confirms' anger to be aimed at vengeance. My account, therefore, provides an attractive positive proposal of anger's nature, as well as support for a diagnosis of why the retributive conception of anger has been so prevalent.

Although those who praise anger take it to be of epistemic value, in being sensitive to reasons for it, and in being able to play knowledge-conducive roles, positive proposals for how anger plays these roles are mostly programmatic. My account of anger's sui generis normative rationality involved arguing for the normative rationality of emotions in general. This is because anger's normative rationality is likely to hinge on features of the kind to which anger belongs: emotion. The account of emotional reason-responsiveness as sui generis made three significant contributions. First, it provided reasons against pursuing the dominant type of approach to normative rationality of emotions, present in the philosophy of emotion literature. These are accounts that take dispositions to comply with norms of rationality to underlie emotional reason-responsiveness. Against this, I argued that emotional reason-responsiveness is best dealt with by taking reasons to be basic. One of the reasons to prefer an account of this sort is that it is better equipped to deal with features of emotional life under conditions of oppression, where it seems necessary to make room of reason-responsiveness as distinct from, and often at the expense of, conformance with rationality requirements. Second, I highlighted how recent work on attentional processing during emotional episodes supports this move to a reasons-basic account. This empirically informed contribution also amounts to a further contribution regarding the nature of anger, namely as

a state that involves multiple types of attentional processes, some of which are not encapsulated from higher-level processing that underlies categorization. Third, I argued that the reason-relation characteristic of emotions is distinctive and irreducible to the epistemic or practical reason-relation. This is a view that matches remarks made by a number of philosophers (de Sousa 1987; Greenspan 1988; Döring 2007) but that has not been cashed out in terms that allow it to be compared to alternative accounts of emotional reason-responsiveness, or that allow its epistemic upshots to be fully appreciated.

The account of sui generis emotional rationality proposed played a role in defending a distinctive epistemic role for anger. Although systematic accounts of the epistemology of anger are scarce in the literature in praise of the emotion, the epistemic role of emotions is a focal topic of debate in contemporary philosophy of emotion. I argued that a crucial part of anger's epistemic value is best cashed out in terms of it playing a justificatory epistemic role. In addition to defending the justificatory epistemic role of emotions against objections present in the literature, I developed and motivated a novel objection, one that arises from taking seriously the epistemic role anger plays under conditions of oppression, as an outlaw emotion. I argued that the justificatory epistemic role of emotions is committed to a defeater clause that risks ruling out the epistemic role of outlaw emotions because these are cases that seem to involve defeaters. I argued that to respond to this objection, an externalist social epistemology is to be preferred. In so doing, I not only highlight an objection that has been neglected in the philosophy of emotion literature, but provided a preferred framework for making sense of the received epistemic commitments in the literature, namely a socially minded externalist one.

The account offered involved taking the potential defeaters in outlaw emotion cases, to be incapable of defeating beliefs based on outlaw emotions, due to their relation to a less reliable epistemic standpoint. That an externalist social epistemology seems to be a promising way of making sense of outlaw emotion cases, suggests that an externalist epistemology might be the preferred framework for making sense of the epistemic role of emotions, more generally. That is to say, I think the proposed framework can be extended to cases where oppressive social relations are not central. For example, take a case of emotional recalcitrance, where one has reason for an emotion but simultaneously holds a belief that conflicts with it. When you have a belief that you should be a doctor but a feeling of uneasiness about becoming one, for example. An externalist epistemology will help make sense of why beliefs based on

recalcitrant emotions are not defeated by conflicting beliefs. In such cases, the conflicting beliefs might be internally justified, but they will not typically be externally justified. Outlaw emotion cases are structurally similar but complicated by the fact that oppressive facts can be self-creating and self-verifying. My account of the sui generis normative rationality of emotions, was also broadly externalist. In so far as I conceived of reasons as facts bearing a basic relation to attitudes, emotional reason-responsiveness will not depend on relations amongst one's beliefs, such as coherence, but on what facts hold. My project has, therefore, suggested that a broadly externalist approach to the epistemic role of emotions seems promising.

I made significant steps towards the provision of an account for how anger plays a distinctive epistemic role under conditions of oppression. In doing so, I have provided theoretical support to the feminist philosophers that highlighted the importance of this epistemic role. Despite having focused on defending the justificatory role of anger in support of evaluative beliefs, my sui generis account of the normative rationality of emotions allows anger to play distinctive *motivational* roles as well. Recall that emotions have been thought not only to provide defeasible justification for evaluative beliefs, but to motivate search for reasons that can lead to evaluative understanding. If emotions enjoy sui generis normative rationality, of the sort outlined in chapter 4, then the evaluative understanding that they can generate through inquiry, much like the evaluative beliefs they can justify, will often be an epistemic achievement made possible only by emotions and their underlying normative rationality.

If emotions provide justification for evaluative beliefs, as I have argued, they seem to act as reasons for these beliefs. How does this square with the notion of reasons I employed in chapter 4, however, where I endeavored to provide an account of emotions themselves as reason-responsive? In chapter 4 I construed reasons as facts which count in favour of certain attitudes. Affective reasons were characterized by strictness and sufficiency. Sufficiency means that they are presumptively sufficient for the justification of emotions. This means that affective reasons are sufficient for the justification of emotions. Strictness means that affective reasons for one emotion cannot be cancelled out by affective reasons for another. Demonstrating both strictness and sufficiency, affective reasons can only be overridden when there is sufficient purely epistemic reason against an affective reason.

Emotions, as reasons for evaluative beliefs, are epistemic reasons. This means that the support emotions provide for evaluative beliefs will be governed by the set of features that characterize epistemic reasons. These included, counting in favour of the truth of a relevant belief, not admitting considerations of the value of holding beliefs, and not demonstrating sufficiency or binarity (chapter 4). If we are committed to the justificatory thesis (JT) that I defended, however, then emotions provide *immediate* defeasible justification to evaluative beliefs. This justificatory thesis can be seen as a claim of presumptive sufficiency regarding the justificatory role of emotions in support of evaluative beliefs. Without JT, emotions would not demonstrate sufficiency in support of evaluative beliefs. This is because emotions are epistemic reasons for evaluative beliefs, and I agreed with others that epistemic reasons do not demonstrate sufficiency. If emotions provide immediate defeasible justification (JT), they seem to be presumptively sufficient for the justification of the beliefs they stand in support of.

Emotions can, however, be cancelled out by other evaluative considerations, because as epistemic reasons, emotions are not strict. Whereas affective reasons (reasons for emotions) are characterized by sufficiency *and* strictness, emotions as reasons for beliefs are epistemic reasons that are only characterized by sufficiency given JT. This means that, emotions that act as reasons for beliefs can be cancelled out by other evaluative considerations. This means that evaluative, as well as epistemic, considerations can jointly provide sufficient reason to undermine an emotion acting as a reason for an evaluative belief. This differs from the case of reasons for emotions, where there must be sufficient purely *epistemic* reason to undermine an affective reason. Reasons for emotions (affective reasons), are therefore distinctive from emotions as reasons in the ways just highlighted.

In construing reasons as facts in chapter 4, however, it may be claimed that I have ruled emotions out from themselves being reasons. After all, emotions are more fit for the label of 'reason' if we construe reasons as mental states, as opposed to facts. I don't think this observation is, in the end, a serious one for me. First, the fact of my being angry might be thought to be the fact that speaks in favour of my associated evaluative belief. In which case these constitute cases where the fact of my being in a given mental state is the reason, construed as a fact. Alternatively, we might offer an alternative way of being a reason, that co-exists with reasons generally being facts. However, the debate over whether emotions can constitute reason, construed as facts, is not one I am going, or need, to settle here. I was

concerned to defend the thesis that emotions themselves can play justificatory roles in support of evaluative beliefs. Whether or not my defense of that thesis is successful depends on the claim that emotions are a source of normative support for evaluative beliefs. Whether we call them 'reasons' – as opposed to, say warrants, or epistemic bases – for belief, once this is granted, seems to be more of a terminological, than a substantive, issue.

6.1 Varieties of Irrationality

My account has helped tease apart distinct ways in which an angry agent can be charged with irrationality. These are:

1. Arationality: when anger is an impulse or physical disturbance outside the realm of reason.
2. Instrumental irrationality:
    a. Undermining the aim associated with one's anger
    b. Undermining one's other aim(s), or one's all-things-considered aim

3. Normative irrationality:
    a. Lacking reasons for one's anger, such that one's anger is inappropriate
    b. Failing to respond to reasons for anger
    c. Violating norms of rationality

4. Epistemic Irrationality:
    a. Unreliable at generating true beliefs about objects of anger
    b. Biasing beliefs unrelated to the object of anger

My account of anger's nature uncontroversially excludes 1, as emotions are ubiquitously taken to be intentional states with some sort of evaluative component. Chapter 3 targeted both 2a and 2b. I have raised the bar for those who wish to condemn anger on a charge of instrumentally irrationality. For anger to be considered typically instrumentally irrational, my opponent will need to provide a story for why cases of non-retributive anger are prevalent in both personal and political life. They will need to provide a story that can account for the empirical work I have called attention to, in chapters 2 and 3. Merely claiming that empirical

work in favour of anger's constructive effects in situations of conflict is unrepresentative of anger, will sit uncomfortably with the evidence levied in support of anger being a heterogeneously constructed kind, canvassed in chapter 2. I have not denied that anger can be instrumentally irrational, in the manners specified in 2a and 2b, but I have denied that instrumental irrationality is prototypical of anger. For an opponent to argue otherwise, merely highlighting the negative effects that anger can cause, will be insufficient. They will presumably have to provide an argument against my positive account, or provide a competing alternative story that makes sense of much of empirical work I have brought to light.

Once arationality is off the table, anger is seen as minimally reason-responsive. The charge of irrationality, in so far as normative irrationality is concerned, then, comes in three forms. First, anger is irrational when it is inappropriate, that is, when there are no reasons for it. Second, failures to feel anger might be rationally condemnable. Third anger may cause the agent to violate norms of rationality. The first sense of irrationality, is an important one that I do not dispute. That anger can be assessed for appropriateness is a fact that this thesis has provided support for and developed. Significantly, I have argued that assessments of anger's instrumental rationality should be kept apart from assessments of anger's normative rationality, and I have put forward an account of emotional reason-responsiveness. We also saw that anger is particularly instrumentally rational when it is perceived to be normatively rational. Someone wishing to condemn anger on the charge of irrationality of variety 3a is often engaged in making the claim that occasions of inappropriate are far too common. This charge relates to the condemnation of anger for reasons of unreliability, which I deal with below.

The second variety of normative irrationality involves failures to feel anger. This is not a typical charge made against anger's rationality. Those who condemn anger, condemn its occurrences rather an its omissions. My project has carved out this neglected sense of rational failure, which speaks to the force of anger as a reason-responsive attitude. Most charges against anger's rationality are of the third kind, that anger violates norms of rationality.

I do not think first-hand experience lends support to such a view. Although anger can distract us from our work and other activities, and can therefore violate enkratic and coherence norms, as well as bias our future interactions with the object of our anger, we do not typically

think of ourselves as incapable of pursuing our goals when we respond to a reason for anger. Sure, some forms of intense rage compromise our judgement and we will need to 'cool down' before a regular task, or interaction, can be subsumed. But this is one extreme form of anger (for which there might be good reason). To claim that it is *typical* of anger that it undermines our capacity to act and reason in conformity with rationality norms is much stronger. Stronger still, seems to be the claim that anger violates norms of rationality, *more often* than it responds to reasons for it. When you hear that the political party you support has been involved in a corruption scheme, when someone maliciously undermines your opinion in public, or when a car cuts you off in a manner that puts you both in danger, you have reason for anger. In responding with anger, you are responding to your affective reasons with an appropriate emotion. Your anger does not, however, typically compromise your ability go on with your day, to continue having an intelligent conversation, or to keep driving skillfully. It would perhaps be a sign of some underlying atypicality if this were not the case of a given individual. It would also be maladaptive, in the evolutionary sense, of anger, if it were to represent such a prominent threat to one's broader functioning.

It could be argued that anger's norm violations *outweigh* its reason-responsive role. That is, it could be argued that anger's norm violations are sufficient to condemn the emotion, because they are so problematic. Establishing this, however, will involve making value judgements about when norm violations matter more than reason-responsiveness and vice-versa. Such judgements are likely to be highly dependent on specific circumstances. We will see that the same is true with epistemic rationality.

A move open to an opponent of mine arguing against anger's epistemic rationality, is similarly, to argue that anger's negative effects *outweigh* its positive ones. This might be the case even if anger does not *typically* cause negative epistemic effects, and even if anger does not cause negative effects *more often* than it causes positive ones. Such a move will involve arguing that anger is unreliable at generating true beliefs about its objects (4a), and that it biases the agent in favour of forming false beliefs that are unrelated to the object of anger (4b). I argued against 4a and 4b in chapter 5 under the label of the 'unreliability objection'. I suggested that we have reason to think emotions are reliable at tracking reasons for them, as well as reason to think that the difference between the biasing effects of emotions, as opposed to that of beliefs and perceptions, have been exaggerated. An opponent wishing to condemn anger on these accounts will have to account for the considerations I raised, and the empirical work levied in support of them. Most importantly, though, in so far as my

argument in favour of anger playing strong and distinctive epistemic roles is successful, an opponent will have to provide enough epistemic downsides to anger, to outweigh them. This sets the bar considerably higher that we might have thought it was.

Pursuing a line of argument that attempts to outweigh anger's positive epistemic value with negative considerations, is unattractive for a number of reasons. First, as I have suggested, if my account of anger's positive effects is accepted, the bar will be higher for how much negative epistemic value must be provided to outweigh the positive. Second, pursuing this line of argument will involve engaging in a weighting game that is hard to carry out. How can we determine how often anger's negative epistemic effects outweigh its positive ones? Similarly, how can we compare the times anger picks up on genuine reasons, against the times it results in violations of different norms of rationality? Pursuing such lines of argument will have clear methodological problems given that these conditions are not easily empirically testable. This highlights a further methodological worry. This is the worry that at each instance, the weighting game will involve value judgements about which epistemic good should be favoured. Such value judgements will be context specific, and, therefore, hard to generalize. Many of us would, for example, take anger's epistemic value to outweigh its negatives, in situations where anger was cueing us to reasons we would otherwise not have access to. This might be the case even if one's anger causes one to hold a number of false beliefs, and to violate coherence amongst our beliefs and intentions. In other circumstances, however, such as when one is taking a test, becoming justifiably angry might be an unwelcome distraction from an important task, as well as potentially compromising of one's performance.

Furthermore, we must bear in mind that a main variable on which the context dependence of these value judgements will depend is one's social position. If one is not a member of an oppressed group, then one might rate the epistemic value of outlaw anger far lower than those who stand to benefit most from such emotions. As academic philosophy is predominantly a white, middle-class, male, field, we should be cautious and mindful of the value judgements made by philosophers regarding the relative value of competing epistemic goods.

In this conclusion I have summarized lines of argument an opponent could pursue, in attempting to undermine the rationality of anger. In the thesis I have provided reasons for thinking these lines of argument are unlikely to be successful, and have sought to make it

harder for anger to be condemned as irrational. My account of anger's rationality did not aim to be exhaustive[136]. I focused on the points of disagreement present in the debate between those who have condemned anger, and those who have most forcefully praised it. I cashed these disagreements out in terms of disagreement over whether anger is, in three distinct senses, rational. In generating desiderata for a foundational approach to the anger debate, and in taking steps to meet them, I hope to have strengthened the case in favour of praising anger.

Let me end by saying that it has also been my aim to bring more clearly into view what exactly the challenge to the claim that anger is rational is. To that extent this thesis has also made things easier for an opponent wishing to condemn anger as irrational: the relevant argumentative moves are open to view. My hope, in pursuing a more foundational approach to the debate, that has made questions about the nature and rationality of anger central and significant, is that I have at the least mapped out a more precise terrain on which future debate can unfold.

---

[136] I have not said much about the role of anger in rationalizing action. I did not set out with such high aims as to provide a complete account of anger's rationality. It is worth noting, however, that the account of anger's rationality that I have provided is likely to be a promising starting point for the development of an account of anger as a source of reasons for action. When anger is triggered by an injustice, the fact that anger has intentional content of the sort that represents, in some way, the situation as unjust, allows the emotion to rationalize the action based on this anger. This is because the emotion contains information related to the wrongness of the occurrence, and that, if invoked, would allow an action done on the basis of it to make sense. A popular view is that emotion's justify actions indirectly by first providing immediate justification to evaluative beliefs (Döring 2007). If this is the case, then the defense I provided of a justificatory epistemic role for emotions in chapter 5, might be a necessary step towards building an account of anger that can justify certain actions. The thought would be, that in providing prima facie justification for an evaluative belief, an emotion supplies the first component in a chain of practical reasoning that can lead to a justified normative judgement of how one ought to act. According to Döring (2007), emotions do not justify normative judgements directly, nor do the evaluative judgements they do justify entail normative judgements. Rather, the evaluative judgements immediately justified by the emotions are important, often crucial, initial steps in the motivation and justification of relevant actions. If such a picture is coupled with my sui generis account of emotional reason-responsiveness, then emotions will plausibly be able to be the source of reasons for action that would not have been accessed in the absence of the emotion.

References

Adams, David. 1995. *Psychology for Peace Activists*. New Haven, CT: The Advocate Press.

Adler, Jonathan. 2017. "Epistemological Problems of Testimony." *The Stanford Encyclopedia of Philosophy*.

Aristotle. 2004. *Rhetoric*. trans. Rhys Roberts. Dover Thrift Editions.

Arpaly, Nomy. 2018. "Comments on Emotions, Values, and Agency by Christine Tappolet" *Philosophy and Phenomenological Research* 97 (2): 520–24.

Averill, James. 1983. "Studies on Anger and Aggression: Implications for Theories of Emotion." *American Psychologist* 38 (11): 1145–60.

Bandura, Albert. 2000. "Exercise of Human Agency Through Collective Efficacy." *Current Directions in Psychological Science* 9 (3): 75–78.

Barrett, Lisa Feldman. 2006. "Are Emotions Natural Kinds?" *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 1 (1): 28–58.

Barton, Richard. 1998. "'Zealous Anger' and the Renegotiation of Aristocratic Relationships in Eleventh- and Twelfth-Century France." In *Rosenwein B. (Ed.), Anger's Past: The Social Uses of an Emotion in the Middle Ages*, 153–70. Cornell University Press.

Basant, Rakesh. 2007. "Social, Economic and Educational Conditions of Indian Muslims." *Economic and Political Weekly* 42 (10): 828–32.

Bell, Macalester. 2009. "Anger, Virtue, and Oppression." In *Feminist Ethics and Social and Political Philosophy: Theorizing the Non-Ideal*, edited by Lisa Tessman, 165–183. Springer.

Ben-Ze'ev, Aaron. 2000. *The Subtlety of Emotions*. MIT Press.

Bergmann, Michael. 1997. "Internalism, Externalism and the no-defeater Condition" *Synthese* 110 (3): 399–417.

Blanchette, Isabelle, Sarah Gavigan, and Kathryn Johnston. 2014. "Does Emotion Help or Hinder Reasoning? The Moderating Role of Relevance." *Journal of Experimental Psychology: General* 143 (3): 1049–64.

Bluhm, Robyn. 2006. "Beyond the Basics: The Evolution and Development of Human Emotions." *Canadian Journal of Philosophy* 36 (sup1): 73–94.

Bodenhausen, Galen, Lori Sheppard, and Geoffrey Kramer. 1994. "Negative Affect and Social Perception: The Differential Impact of Anger and Sadness." *European Journal of Social Psychology* 24: 45–62.

Bommarito, Nicolas. 2017. "Virtuous and Vicious Anger." *Journal of Ethics and Social Philosophy* 11 (3): 1–28.

Boyd, Robert, and Peter Richerson. 1985. *Culture and the Evolutionary Process*. The University of Chicago Press.

Brady, Michael. 2010. "Virtue, Emotion and Attention." *Metaphilosophy* 41 (1–2).

———. 2011. "Emotions, Perceptions, and Reasons." In *Morality and the Emotions*. Oxford University Press.

———. 2013. *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford University Press.

———. 2014. *Emotion, Attention, and the Nature of Value*. Oxford University Press.

Briggs, Jean. 1971. *Never in Anger*. Harvard University Press.

Broome, John. 1999. "Normative Requirements." *Ratio* 12 (4): 398–419.

Brosnan, Sarah. 2006. "Nonhuman Species' Reactions to Inequity and Their Implications for Fairness." *Social Justice Research* 19 (2): 153–85.

Brown, Cynthia Stokes. 2002. *Refusing Racism: White Allies and the Struggle for Civil Rights*. The Teaching for Social Justice Series. New York: Teachers College Press.

Burnyeat, M. F. 2002. "Excuses for Madness." *London Review of Books*, October 17, 2002.

Calhoun, Cheshire. 2004. *Setting the Moral Compass: Essays by Women Philosophers*. Studies in

Feminist Philosophy. Oxford ; New York: Oxford University Press.

Carretié, Luis. 2014. "Exogenous (Automatic) Attention to Emotional Stimuli: A Review." *Cognitive, Affective, & Behavioral Neuroscience* 14 (4): 1228–58.

Chakravarti, Sonali. 2014. *Sing the Rage.* The University of Chicago Press.

Chareyron, Loïc, Pamela Banta Lavenex, David Amaral, and Pierre Lavenex. 2011. "Stereological Analysis of the Rat and Monkey Amygdala." *The Journal of Comparative Neurology* 519 (16): 3218–39.

Charland, Louis. 2002. "The Natural Kind Status of Emotion." *British Journal for the Philosophy of Science* 53 (4): 511–37.

Cherry, Myisha, and Owen Flanagan. 2017. "Introduction." In *The Moral Psychology of Anger*, Edited by Myisha Cherry and Owen Flanagan.

Christensen, Niels. 2016. *Aristotle on Anger and Rectificatory Justice.* MPhil thesis, Department of Philosophy, University College London.

Clayson, Peter and Michael Larson. 2013. "Adaptation to Emotional Conflict: Evidence from a Novel Face Emotion Paradigm." Edited by Piia Susanna Astikainen. *PLoS ONE* 8 (9): e75776.

Clore, Gerald, and Karen Gasper. 2000. "Feeling Is Believing: Some Affective Influences on Belief." *Emotions and Beliefs - How Feelings Influence Thoughts*, January.

Cogley, Zac. 2014. *A Study of Virtuous and Vicious Anger.* Oxford University Press. https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199645541.001 .0001/acprof-9780199645541-chapter-10.

Cohen, Jonathan. 2009. "Rationality." In *A Companion to Epistemology*, edited by Jonathan Dancy, Ernest Sosa, and Matthias Steup, 2nd ed., 663–68. Wiley-Blackwell.

Colombetti, Giovanna. 2013. *The Feeling Body: Affective Science Meets the Enactive Mind.* MIT Press.

Cowan, Robert. 2016. "Epistemic Perceptualism and Neo-Sentimentalist Objections." *Canadian Journal of Philosophy* 46 (1): 59–81.

———. 2018. "Epistemic Sentimentalism and Epistemic Reason-Responsiveness." In *Evaluative Perception*, edited by Anna Bergqvist and Robert Cowan. Oxford University Press.

Crisp, Roger. 2000. *Aristotle: Nicomachean Ethics.* Translated by Crisp. Cambridge, U.K. ; New York: Cambridge University Press.

Damasio, Antonio. 2005. *Descartes' Error: Emotion, Reason and the Human Brain.* Penguin Random House.

Danielsson, Sven, and Jonas Olson. 2007. "Brentano and the Buck-Passers." *Mind* 116 (463): 511–22.

D'Arms, Justin, and Daniel Jacobson. 2000. "The Moralistic Fallacy: On the 'Appropriateness' of Emotions." *Philosophy and Phenomenological Research* 61 (1): 65.

Darwall, Stephen. 2013. *Ressentiment and Second-Personal Resentment.* Oxford University Press.

Darwin, Charles. 1872. *The Expression of the Emotions in Man and Animals.* The Expression of the Emotions in Man and Animals. London, England: John Murray.

de Sousa, Ronald. 1987. *The Rationality of Emotion.* Cambridge, Mass: MIT Press.

———. 2007. "Truth, Authenticity, and Rationality." Dialectica 61 (3): 323–45.

DeLancey, Craig. 2002. *Passionate Engines: What Emotions Reveal about the Mind and Artificial Intelligence.* Oxford University Press.

Deonna, Julien, and Fabrice Teroni. 2012a. *The Emotions: A Philosophical Introduction.* London ; New York: Routledge.

———. 2012b. "From Justified Emotions to Justified Evaluative Judgements." *Dialogue* 51 (01): 55–77.

———. 2015. "Emotions as Attitudes." *Dialectica* 69 (3): 293–311.

Deonna, Julien, Christine Tappolet, and Fabrice Teroni. 2015. "Emotion, Philosophical

Issues about: Emotion, Philosophical Issues About." *Wiley Interdisciplinary Reviews: Cognitive Science* 6 (3): 193–207.

Döring, Sabine. 2003. "Explaining Action by Emotion." *The Philosophical Quarterly* 53 (211): 214–30.

———. 2007. "Seeing What to Do: Affective Perception and Rational Motivation." *Dialectica* 61 (3): 363–94.

———. 2009. "The Logic of Emotional Experience: Noninferentiality and the Problem of Conflict Without Contradiction." *Emotion Review* 1 (3): 240–47.

Döring, Sabine, and Anika Lutz. 2015. "Beyond Perceptualism: Introduction to the Special Issue." *Dialectica* 69 (3): 259–70.

Dotson, Kristie. 2012. "A Cautionary Tale: On Limiting Epistemic Oppression." *Frontiers: A Journal of Women Studies* 33 (1): 24–47.

Douglass, Frederick. 2009. *Narrative of the Life of Frederick Douglass An American Slave, Written by Himself.* Harvard University Press.

Du Bois, W.E.B. 1996. *The Souls of Black Folk.* Penguin Classics.

Dunn, Jennifer, and Maurice E. Schweitzer. 2005. "Feeling and Believing: The Influence of Emotion on Trust." *Journal of Personality and Social Psychology* 88

Ekman, Paul. 1982. *Emotion in the Human Face.* 2nd ed. Cambridge University Press.

———. 1999. "Basic Emotions." In *Handbook of Cognition and Emotion*, ed. P. Ekman and R. Davidson, 45–60. John Wiley & Sons, Ltd.

Ekman, Paul, and Wallace Friesen. 1971. "Constants across Cultures in the Face and Emotion." *Journal of Personality and Social Psychology* 17 (2): 124–29.

Etkin, Amit, Tobias Egner, Daniel Peraza, Eric Kandel, and Joy Hirsch. 2006. "Resolving Emotional Conflict: A Role for the Rostral Anterior Cingulate Cortex in Modulating Activity in the Amygdala." *Neuron* 51 (6): 871–82.

Fanon, Frantz. 2008. *Black Skin, White Masks.* Grove Press.

Faraci, David. 2018. "We Have No Reason to Think There Are No Reasons for Affective Attitudes." *Mind.*

Faucher, Luc, and Christine Tappolet. 2002. "Fear and the Focus of Attention." *Consciousness & Emotion* 3 (2): 105–44.

———. 2006. "Introduction: Modularity and the Nature of Emotions." *Canadian Journal of Philosophy* 36 (suplement 1). https://philarchive.org.

Fazekas, Peter, and Bence Nanay. 2018. "Attention Is Amplification, Not Selection." *The British Journal for the Philosophy of Science*, October.

Fessler, Daniel, and Kevin Haley. 2003. "The Strategy of Affect: Emotions in Human Cooperation." In *Genetic and Cultural Evolution of Cooperation*, 7–36. Dahlem Workshop Report. Cambridge, MA, US: MIT Press.

Finucane, Melissa, Paul Slovic, Stephen Johnson, and Ali Alhakami. 2000. "The Affect Heuristic in Judgments of Risks and Benefits." *Journal of Behavioral Decision Making.*

Fodor, Jerry. 1983. *The Modularity of Mind.* MIT Press.

Frank, Robert. 1988. *Passions within Reason: The Strategic Role of the Emotions.* New York, NY: Norton.

Fredrickson, Barbara, and Christine Branigan. 2005. "Positive Emotions Broaden the Scope of Attention and Thought-Action Repertoires." *Cognition & Emotion* 19 (3): 313–32.

———. 2000. "Oppression." In *Encyclopedia of Feminist Theories*, edited by Lorraine Code, 370. London & New York: Routledge.

Freedman, Paul. 1998. "Peasant Anger in the Late Middle Ages." In *Anger's Past: The Social Uses of an Emotion in the Middle Ages*, Barbara H. Rosenwein, 171–88. Cornell University Press.

Fricker, Miranda. 1991. "Reason and Emotion." *Radical Philosophy* 57 (Spring): 14–19.

———. 1999. "Epistemic Oppression and Epistemic Privilege." *Canadian Journal of Philosophy* 29 (Supplement): 191–210.

———. 2007. *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford University Press.

Friedman, Marilyn. 1986. "Autonomy and the Split-Level Self." *The Southern Journal of Philosophy* 24 (1): 19–35.

Frijda, Nico. 2005. "Emotion Experience: Cognition and Emotion: Vol 19, No 4." 2005.

Frye, Marilyn. 1983. *The Politics of Reality.* Crossing Press.

Godfrey-Smith, Peter. 2008. "Varieties of Population Structure and the Levels of Selection." *The British Journal for the Philosophy of Science* 59 (1): 25–50.

Goldie, Peter. 2002. *The Emotions: A Philosophical Exploration.* Oxford University Press.

———. 2004a. "Emotion, Feeling, and Knowledge of the World." In *Thinking About Feeling: Contemporary Philosophers on Emotions*, edited by Robert C. Solomon. Oxford University Press.

———. 2004b. "Emotion, Reason, and Virtue." In *Emotion, Evolution, and Rationality*, edited by Dylan Evans and Pierre Cruse, 250–66. Oxford University Press.

———. 2006. "Review: Gut Reactions: A Perceptual Theory of Emotion." *Mind* 115 (458): 453–57.

Goldman, Alvin. 1986. *Epistemology and Cognition.* Epistemology and Cognition. Cambridge, MA, US: Harvard University Press.

Gordon, Robert. 1987. *The Structure of Emotions: Investigations in Cognitive Philosophy.* Cambridge University Press.

Greenspan, Patricia. 1988. *Emotions and Reasons: An Enquiry Into Emotional Justification.* Routledge.

Griffiths, Paul. 1997. *What Emotions Really Are: The Problem of Psychological Categories.* The University of Chicago Press.

———. 2004. "Emotions as Natural and Normative Kinds." *Philosophy of Science* 71 (5): 901–11.

Griffiths, Paul, and Karola Stotz. 2013. *Genetics and Philosophy: An Introduction.* Cambridge: Cambridge University Press.

Grimm, Stephen. 2006. "Is Understanding a Species of Knowledge?" *British Journal for the Philosophy of Science* 57 (3): 515–535.

Hacking, Ian. 1999. *The Social Construction of What?* Harvard University Press.

Hajcak, Greg, Annmarie MacNamara, and Doreen M. Olvet. 2010. "Event-Related Potentials, Emotion, and Emotion Regulation: An Integrative Review." *Developmental Neuropsychology* 35 (2): 129–55.

Halperin, Eran. 2008. "Group-Based Hatred in Intractable Conflict in Israel." *The Journal of Conflict Resolution* 52 (5): 713–36.

Harding, Sandra. 2004. "A Socially Relevant Philosophy of Science? Resources from Standpoint Theory's Controversiality." *Hypatia* 19 (1): 25–47.

Hartsock, Nancy. 1983. "The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism." In *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, edited by Sandra Harding and Merrill B. Hintikka, 283–310. Synthese Library. Dordrecht: Springer Netherlands.

Haslanger, Sally. 2011. "Ideology, Generics, and Common Ground." In *Feminist Metaphysics*, edited by Charlotte Witt, 179–207. Dordrecht: Springer Netherlands.

———. 2012. *"But Mom, Crop-Tops Are Cute!": Social Knowledge, Social Structure, and Ideology Critique.* Oxford University Press.

Helm, Bennett W. 2001. *Emotional Reason: Deliberation, Motivation, and the Nature of Value.* Cambridge University Press.

Hessel, Stephane. 2012. *Indignez-Vous!* Montpellier: Indigene éditions.

Hieronymi, Pamela. 2001. "Articulating an Uncompromising Forgiveness." *Philosophy and Phenomenological Research* 62 (3): 529–55.

Holmes, Mary. 2004. "The Importance of Being Angry: Anger in Political Life." *European Journal of Social Theory* 7 (2): 123–32.

hooks, bell. 1995. *Killing Rage: Ending Racism*. New York: H. Holt and Co.

Hughlings-Jackson, J. 1879. "On Affectations of Speech From Disease of The Brain." *Brain* 2 (3): 323–56.

Izard, Carroll. 1977. *Human Emotions*. Emotions, Personality, and Psychotherapy. Springer US.

———. 1994. "Innate and Universal Facial Expressions: Evidence from Developmental and Cross-Cultural Research." *Psychological Bulletin* 115 (2): 288–99.

Jablonka, Eva, and Marion Lamb. 2006. *Evolution in Four Dimensions*. MIT Press.

Jaggar, Alison. 1989. "Love and Knowledge: Emotion in Feminist Epistemology." *Inquiry* 32 (2): 151–76.

James, Susan. 1997. *Passion and Action: The Emotions in Seventeenth-Century Philosophy*. Oxford : New York: Clarendon Press ; Oxford University Press.

Jasper, James M. 2014. "Constructing Indignation: Anger Dynamics in Protest Movements." *Emotion Review* 6 (3): 208–13.

Johnston, Mark. 2001. "The Authority of Affect*." *Philosophy and Phenomenological Research* 63 (1): 181–214.

Jones, Karen. 2003. "Emotion, Weakness of Will, and the Normative Conception of Agency." *Royal Institute of Philosophy Supplement* 52 (March): 181–200.

Kaplow, Susi. 1971. "Getting Angry." In *Feminism, Edited by Anne Koedt, Ellen Levine, and Anita Rapone*, 36–47. New York: Quadrangle, 1973.

Keltner, Dacher, and K. Edwards. 1993. "Beyond Simple Pessimism: Effects of Sadness and Anger on Social Perception." *Journal of Personality and Social Psychology* 64 (5): 740–52.

Keltner, Dacher, and Jonathan Haidt. 1999. "Social Functions of Emotions at Four Levels of Analysis." *Cognition & Emotion* 13 (5): 505–21.

Kennedy, George. 2006. *Aristotle: On Rhetoric: A Theory of Civic Discourse*. 2nd ed. Oxford University Press.

Kenny, Anthony. 1963. *Action, Emotion And Will*. Ny: Humanities Press.

Kitcher, Philip. 1994. "Contrasting Conceptions of Social Epistemology." In *Socializing Epistemology: The Social Dimensions of Knowledge*, ed. Frederick F. Shmitt. Lanham, Md.: Rowman & Littlefield.

Klimecki, Olga, David Sander, and Patrik Vuilleumier. 2018. "Distinct Brain Areas Involved in Anger versus Punishment during Social Interactions." *Scientific Reports* 8 (1): 10556.

Kolnai, Aurel. 1998. "The Standard Modes of Aversion: Fear, Disgust and Hatred." *Mind* 107 (427): 581–96.

Kolodny, Niko. 2005. "Why Be Rational?" *Mind* 114 (455): 509–63.

Konstan, David. 2007. *The Emotions of the Ancient Greeks: Studies in Aristotle and Classical Literature*. Edição: 1. Toronto: University of Toronto Press, Scholarly Publishing Division.

Korsgaard, Christine. 1996. *The Sources of Normativity*. Vol. 110. Cambridge University Press.

Kurth, Charlie. 2018. "Are Emotions Psychological Constructions?" *Philosophy of Science*.

Kurth, Charlie, Haley Crosby, and Jack Basse. 2018. "Are Emotions Perceptions of Value ? A Review Essay of Christine Tappolet's Emotions, Values, and Agency." *Philosophical Psychology* 31 (4): 483–499.

Langton, Rae. 2009. *Sexual Solipsism: Philosophical Essays on Pornography and Objectification*. Oxford University Press.

Lazarus, Richard. 1991. *Emotion and Adaptation*. Emotion and Adaptation. New York, NY, US: Oxford University Press.

Leboeuf, Celine. 2017. "Anger as a Political Emotion: A Phenomenological Perspective." In *The Moral Psychology of Anger*, edited by Myisha Cherry and Owen Flanagan, 15–30.

Lepoutre, Maxime. 2018. "Rage Inside the Machine: Defending the Place of Anger in Democratic Speech." *Politics, Philosophy and Economics*.

Lerner, Jennifer, and Dacher Keltner. 2001. "Fear, Anger, and Risk." *Journal of Personality and Social Psychology* 81 (1): 146–59.

Lerner, Jennifer, Julie Goldberg, and Philip Tetlock. 1998. "Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility." *Personality and Social Psychology Bulletin* 24 (6): 563–74.

Lerner, Jennifer, and Larissa Tiedens. 2006. "Portrait of the Angry Decision Maker: How Appraisal Tendencies Shape Anger's Influence on Cognition." *Journal of Behavioral Decision Making* 19 (2): 115–37.

Levy, Robert. 1973. *Tahitians: Mind and Experience in the Society Islands*. The University of Chicago Press.

Lorde, Audre. 1997. "The Uses of Anger." *Women's Studies Quarterly*. https://academicworks.cuny.edu/wsq/509.

Lugones, Maria. 2003. *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions -*. Rowman & Littlefield Publishers.

Lutz, Catherine. 1988. *Unnatural Emotions: Everyday Sentiments on a Micronesian Atoll & Their Challenge to Western Theory*. Chicago: University of Chicago Press.

———. 2002. "Emotions and Feminist Theories." In *Querelles: Jahrbuch Für Frauenforschung 2002*, edited by Ingrid Kasten, Gesa Stedman, and Margarete Zimmermann, 104–21. Stuttgart: J.B. Metzler.

Mackinnon, Catharine. 1987. *Feminism Unmodified Discourses on Life and Law*.

MacLachlan, Alice. 2010. "Resentment and Moral Judgment in Smith and Butler." *The Adam Smith Review* 5: 161–177.

Mallon, Ron. 2007. "A Field Guide to Social Construction." *Philosophy Compass* 2 (1): 93–108.

Mallon, Ron, and Stephen Stich. 2000. "The Odd Couple: The Compatibility of Social Construction and Evolutionary Psychology." *Philosophy of Science* 67 (1): 133–54.

Marques, Teresa. 2015. "Construção Social." In *Compêndio Em Linha de Problemas de Filosofia Analítica*, edited by João Branquinho and Ricardo Santos, 1–23.

Mathews, Andrew, and Colin MacLeod. 1985. "Selective Processing of Threat Cues in Anxiety States." *Behaviour Research and Therapy* 23 (5): 563–69.

McDonald, Melissa, Roni Porat, Ayala Yarkoney, Michal Tagar, Sasha Kimel, Tamar Saguy, and Eran Halperin. 2017. "Intergroup Emotional Similarity Reduces Dehumanization and Promotes Conciliatory Attitudes in Prolonged Conflict." *Group Processes & Intergroup Relations* 20 (1): 125–36.

McDougall, William. 1926. *An Introduction to Social Psychology*. Boston: Luce.

McHugh, Conor, and Jonathan Way. 2018. "What Is Good Reasoning?" *Philosophy and Phenomenological Research* 96 (1): 153–74.

McIntyre, A., 1990, "Is Akratic Action Always Irrational?," in *Identity, Character, and Morality*, O. Flanagan and A. Rorty (eds.), Cambridge, MA: MIT Press, pp. 379-400

Miceli, Maria, and Cristiano Castelfranchi. 2017. "Anger and Its Cousins." *Emotion Review* 11 (1): 13–26.

Michaelian, Kourken. 2008. "Privileged Standpoints/Reliable Processes." *Hypatia* 23 (1): 65–98.

Miller, David. 2017. "Justice." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2017. Metaphysics Research Lab, Stanford University.

Mitchell, Jonathan. 2017. "The Epistemology of Emotional Experience." *Dialectica* 71 (1): 57–84.

Moretti, Luca, and Tommaso Piazza. 2018. "Defeaters in Current Epistemology: Introduction to the Special Issue." *Synthese* 195 (7): 2845–54.

Mulligan, Kevin. 2007. "Intentionality, Knowledge and Formal Objects." *Disputatio* 2 (23): 1–24.

Narayan, Uma. 1988. "Working Together Across Difference: Some Considerations on Emotions and Political Practice." *Hypatia* 3 (2): 31–48.

Nussbaum, Martha. 2004. "Emotions as Judgments of Value and Importance." In *Thinking about Feeling: Contemporary Philosophers on Emotions*, 183–99. Series in Affective Science. New York, NY, US: Oxford University Press.

———. 2015. "Transitional Anger." *Journal of the American Philosophical Association* 1 (01): 41–56.

———. 2016. *Anger and Forgiveness: Resentment, Generosity, and Justice*. Oxford, New York: Oxford University Press.

Okasha, Samir. 2001. "Why Won't the Group Selection Controversy Go Away?" *The British Journal for the Philosophy of Science* 52 (1): 25–50.

Okasha, Samir, and Chris Paternotte. 2012. "Group Adaptation, Formal Darwinism and Contextual Analysis," 25(6): 1127–1139.

Ortells, Juan, Concepcion Vellido, Maria Teresa Daza, and Carmen Noguera. 2006. "Semantic Priming Effects with and without Perceptual Awareness." *Psicologica: International Journal of Methodology and Experimental Psychology* 27 (2): 225–42.

Ortony, Andrew, and Terence Turner. 1990. "What's Basic about Basic Emotions?" *Psychological Review* 97 (3): 315–31.

Panksepp, Jaak. 1988. "Brain Emotional Circuits and Psychopathologies." In *Emotions and Psychopathology*, 37–76. New York, NY, US: Plenum Press.

Pelser, Adam C. 2014. *Emotion, Evaluative Perception, and Epistemic Justification*. Oxford University Press.

Pessoa, Luiz. 2010. "Emotion and Cognition and the Amygdala: From 'What Is It?' To 'What's to Be Done?'" *Neuropsychologia* 48 (12): 3416–29.

———. 2014. "Understanding Brain Networks and Brain Organization." *Physics of Life Reviews* 11 (3): 400–435.

Pettigrove, Glen. 2012. "Meekness and 'Moral' Anger." *Ethics* 122 (2): 341–70.

Plamper, Jan. 2017. *The History of Emotions: An Introduction*. Emotions in History. Oxford, New York: Oxford University Press.

Plantinga, Alvin. 1993. *Bayesian Coherentism and Rationality*. Oxford University Press.

Plutchik, Robert. 1962. *The Emotions: Facts, Theories and a New Model*. The Emotions: Facts, Theories and a New Model. New York, NY, US: Crown Publishing Group/Random House.

Pollock, John L. 1986. *Contemporary Theories of Knowledge*. Savage, MD: Rowman and Littlefield.

Prinz, Jesse. 2006. *Gut Reactions: A Perceptual Theory of Emotion*. Philosophy of Mind. Oxford, New York: Oxford University Press.

Pugmire, David. 2005. *Sound Sentiments: Integrity in the Emotions*. Oxford, New York: Oxford University Press.

Ratcliffe, Matthew. 2005. "William James on Emotion and Intentionality." *International Journal of Philosophical Studies* 13 (2): 179–202.

Raz, Joseph. 1999. *Engaging Reason: On the Theory of Value and Action*. Oxford ; New York: Oxford University Press.

———. 2011. "Reasons: Practical and Adaptive." In *From Normativity to Responsibility*. Oxford University Press.

Risen, Clay. 2009. *A Nation on Fire: America in the Wake of the King Assassination*. Hoboken, N.J.: John Wiley & Sons.

Roberts, Robert. 2003. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge University Press.

Robinson, Jenefer. 2005. *Deeper than Reason: Emotion and Its Role in Literature, Music, and Art*. Oxford University Press.

Rosen, Gideon. n.d. "Rage Against The Machine: Anger as a Political Emotion." *In Progress*.

Rosenmann, Marcos. 2012. *Los indignados. El rescate de la política*.

Rosenwein, Barbara. ed. 1998. *Anger's Past: The Social Uses of an Emotion in the Middle Ages*. Ithaca, NY: Cornell University Press.

Russell, James. 1994. "Is There Universal Recognition of Emotion from Facial Expression? A Review of the Cross-Cultural Studies." *Psychological Bulletin* 115 (1): 102–41.

———. 2003. "Core Affect and the Psychological Construction of Emotion." *Psychological Review* 110 (1): 145–172.

Salmela, Mikko. 2011. "Can Emotion Be Modelled on Perception?: Can Emotion Be Modelled on Perception?" *Dialectica* 65 (1): 1–29.

———. 2014. *True Emotions*. Consciousness & Emotion Book Series, v. 9. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Śāntideva. 1997. *The Way of the Bodhisattva : A Translation of the Bodhicharyāvatāra* Padmakara Translation Group, Boston: Shambhala

Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, Mass: Belknap Press of Harvard University Press.

———. 2014. *Being Realistic about Reasons*. Oxford: Oxford University Press.

Scarantino, Andrea. 2010. "Insights and Blindspots of the Cognitivist Theory of Emotions." *The British Journal for the Philosophy of Science* 61 (4): 729–68.

Scarantino, Andrea, and Paul Grifftiths. 2011. "Don't Give Up on Basic Emotions." *Emotion Review* 3 (4): 444–454.

Scheepers, Daan, Russell Spears, Bertjan Doosje, and Antony Manstead. 2006. "Diversity in In-Group Bias: Structural Factors, Situational Features, and Social Functions." *Journal of Personality and Social Psychology* 90 (6): 944–60.

Scherer, Klaus. 2005. "Appraisal Theory." In *Handbook of Cognition and Emotion*, 637–63. John Wiley & Sons, Ltd.

Schmitter, Amy. 2016. "17th and 18th Century Theories of Emotions." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2016. Metaphysics Research Lab, Stanford University.

Schroeder, Tim. 2017. "Desire." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2017. Metaphysics Research Lab, Stanford University.

Schroer, Robert. 2008. "Memory Foundationalism and the Problem of Forgotten Carelessness." *Pacific Philosophical Quarterly* 89 (March): 74–85.

Scott, Jill. 2010. "Anger without Emotion: Revenge from the Iliad to Kill Bill." In *A Poetics of Forgiveness: Cultural Responses to Loss and Wrongdoing*, edited by Jill Scott, 25–43. New York: Palgrave Macmillan US.

Sell, Aaron, John Tooby, and Leda Cosmides. 2009. "Formidability and the Logic of Human Anger." *Proceedings of the National Academy of Sciences* 106 (35): 15073–78.

Seneca, Lucius Annaeus. 2010. "Anger, Mercy, Revenge." In *The Complete Works of Lucius Annaeus Seneca*. Chi- cago: University of Chicago Press.

Senor, Thomas. 1996. "The Prima/Ultima Facie Justification Distinction in Epistemology." *Philosophy and Phenomenological Research* 56 (3): 551–566.

———. 2014. "Epistemological Problems of Memory." *The Stanford Encyclopedia of Philosophy*.

Shuman, Eric, Eran Halperin, and Michal Reifen Tagar. 2018. "Anger as a Catalyst for

Change? Incremental Beliefs and Anger's Constructive Effects in Conflict." *Group Processes & Intergroup Relations* 21 (7): 1092–1106.

Siegel, Susanna. 2017. *The Rationality of Perception*. Oxford, New York: Oxford University Press.

Silvermint, Daniel. 2017. "Rage and Virtuous Resistance." *Journal of Political Philosophy* 25 (4): 461–86.

Skorupski, John. 2010. *The Domain of Reasons*. Oxford University Press.

Smith, Adam. 1976. *The Theory of Moral Sentiments*. D.D. Raphael and A.L. Macfie. Oxford: Clarendon Press: Liberty Press Imprint.

Smith, Michael. 1994. *The Moral Problem*. Blackwell.

Sober, Elliott. 1991. "Models of Cultural Evolution." In *Trees of Life: Essays in Philosophy of Biology*, edited by Paul Griffiths, 477–492. Australasian Studies in History and Philosophy of Science. Dordrecht: Kluwer.

Solomon, Robert. 1976. *The Passions*. University of Notre Dame Press.

———. 1984. "Getting Angry: The Jamesian Theory of Emotion in Anthropology." In *Not Passion's Slave*. Oxford University Press 2003.

———. 2003. *Not Passion's Slave*. Oxford University Press.

Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford, New York: Oxford University Press.

Spelman, Elizabeth. 1989. "Anger and Insubordination." In *Women, Knowledge, and Reality*, ed. Ann Garry and Marilyn Pearsall. Unwin Hyman Ltd,.

Spring, Victoria, Daryl Cameron, and Mina Cikara. 2018. "The Upside of Outrage." *Trends in Cognitive Sciences* 22 (12): 1067–69.

Srinivasan, Amia. 2018. "The Aptness of Anger: APTNESS OF ANGER." *Journal of Political Philosophy* 26 (2): 123–44.

———. in progress. "Radical Externalism." https://philarchive.org/archive/SRIRE

Stearns, Carol, and Peter Stearns. 1986. *Anger: The Struggle for Emotional Control in America's History*. University of Chicago Press.

Sterelny, Kim. 2006. "The Evolution and Evolvability of Culture," no. 21: 137-165.

———. 2016. "Cooperation, Culture, and Conflict." *The British Journal for the Philosophy of Science* 67 (1): 31–58.

Stocker, Michael, and Elizabeth Hegeman. 1996. *Cambridge Studies in Philosophy: Valuing Emotions*. Cambridge University Press.

Strawson, Peter F. 1962. "Freedom and Resentment." In *Proceedings of the British Academy, Volume 48: 1962*, edited by Gary Watson, 1–25. Oup Oxford.

Stroop, Ridley. 1935. "Studies of Interference in Serial Verbal Reactions." Journal of Experimental Psychology, no. 18: 643–62.

Suchak, Malini, Timothy M. Eppley, Matthew W. Campbell, Rebecca A. Feldman, Luke F. Quarles, and Frans B. M. de Waal. 2016. "How Chimpanzees Cooperate in a Competitive World." *Proceedings of the National Academy of Sciences* 113 (36): 10215–20.

Swaine, Lucas A. 1996. "Blameless, Constructive, and Political Anger." *Journal for the Theory of Social Behaviour* 26 (3): 257–74.

Tagar, Michal Reifen, Christopher M. Federico, and Eran Halperin. 2011. "The Positive Effect of Negative Emotions in Protracted Conflict: The Case of Anger."

Tappolet, Christine. 1997. "The Irrationality of Emotions." *Université de Montréal*, no. Conference Paper.

———. 2000. *Emotions et valeurs*. 1re éd. Philosophie morale. Paris: Presses universitaires de France.

———. 2005. "Ambivalent Emotions and the Perceptual Account of Emotions." *Analysis* 65 (287): 229–33.

———. 2009. "Emotion, Motivation, and Action: The Case of Fear." *The Oxford Handbook*

of Philosophy of Emotion, December.

———. 2014. "Emotions, Reasons, and Autonomy." In *Autonomy, Oppression and Gender*, edited by Andrea Veltman and Mark C. Piper, 163–180. Oxford University Press.

———. 2016. *Emotions, Values, and Agency*. Oxford: Oxford University Press.

———. 2018a. "Précis of Emotions, Values, and Agency." *Philosophy and Phenomenological Research* 97 (2): 494–99.

———. 2018b. "Reasons and Emotions." In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star, 892–904.

———. 2018c. "Replies." *Philosophy and Phenomenological Research* 97 (2): 525–37.

Tausch, Nicole, Julia C. Becker, Russell Spears, Oliver Christ, Rim Saab, Purnima Singh, and Roomana N. Siddiqui. 2011. "Explaining Radical Group Behavior: Developing Emotion and Efficacy Routes to Normative and Nonnormative Collective Action." *Journal of Personality and Social Psychology* 101 (1): 129–48.

Teroni, Fabrice. 2007. "Emotions and Formal Objects." *Dialectica* 61 (3): 395–415.

Thompson, Simon. 2006. "Anger and the Struggle for Justice." In *Emotion, Politics and Society*, edited by Simon Clarke, Paul Hoggett, and Simon Thompson, 123–44. London: Palgrave Macmillan UK.

Tompkins, Silvan. 1963. *Affect, Imagery, Consciousness: II. The Negative Affects*. Affect, Imagery, Consciousness: II. The Negative Affects. Oxford, England: Springer.

Tooby, John, and Leda Cosmides. 2008. "The Evolutionary Psychology of the Emotions and Their Relationship to Internal Regulatory Variables." In *Handbook of Emotions, 3rd Ed*, 114–37. New York, NY, US: The Guilford Press.

Toole, Briana. forthcoming. "From Standpoint Epistemology to Epistemic Oppression." *Hypatia: A Journal of Feminist Philosophy*.

Van Kleef, Gerben A., and Stéphane Côté. 2007. "Expressing Anger in Conflict: When It Helps and When It Hurts." *The Journal of Applied Psychology* 92 (6): 1557–69.

Vos, Bart de, Martijn van Zomeren, Ernestine H. Gordijn, and Tom Postmes. 2013. "The Communication of 'Pure' Group-Based Anger Reduces Tendencies toward Intergroup Conflict Because It Increases out-Group Empathy." *Personality & Social Psychology Bulletin* 39 (8): 1043–52.

———. 2016. "When Does the Communication of Group-Based Anger Increase Outgroup Empathy in Intergroup Conflict? The Role of Perceived Procedural Unfairness and Outgroup Consensus." *Group Processes & Intergroup Relations* 21 (4): 533–48.

Waal, Frans de. 1992. "The Chimpanzee's Sense of Social Regularity and Its Relation to the Human Sense of Justice." In *The Sense of Justice: Biological Foundations of Law*, 241–55. Sage Focus Editions, Vol. 136. Thousand Oaks, CA, US: Sage Publications, Inc.

———. 1997. *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Harvard University Press.

———. 2014. "Natural Normativity: The 'Is' and 'Ought' of Animal Behavior." *Behaviour* 151 (2–3): 185–204.

Walker, Margaret. 2006. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. Cambridge University Press.

Williams, Riccardo. 2017. "Anger as a Basic Emotion and Its Role in Personality Building and Pathological Growth: The Neuroscientific, Developmental and Clinical Perspectives." *Frontiers in Psychology* 8 (November).

Wlodarczyk, Anna, Nekane Basabe, Darío Páez, and Larraitz Zumeta. 2017. "Hope and Anger as Mediators Between Collective Action Frames and Participation in Collective Mobilization: The Case of 15-M." *Journal of Social and Political Psychology* 5 (1): 200-223–223.

Worsnip, Alex. 2016. "Moral Reasons, Epistemic Reasons and Rationality." *The Philosophical*

Quarterly 66 (263): 341–61.

Yancy, George. 2008. "Elevators, Social Spaces and Racism: A Philosophical Analysis." *Philosophy & Social Criticism* 34 (8): 843–76.

Yoo, Seung Hee, Margaret Clark, Edward Lemay, Peter Salovey, and Joan Monin. 2011. "Responding to Partners' Expression of Anger: The Role of Communal Motivation." *Personality & Social Psychology Bulletin* 37 (2): 229–41.

Young, Iris M. 1988. "Five Faces of Oppression." *Philosophical Forum* 19 (4): 270.

Zomeren, Martijn van, Colin Leach, and Russell Spears. 2012. "Protesters as 'Passionate Economists': A Dynamic Dual Pathway Model of Approach Coping with Collective Disadvantage." *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc* 16 (2): 180–99.

Zomeren, Martijn van, Russell Spears, Agneta Fischer, and Colin Leach. 2004. "Put Your Money Where Your Mouth Is! Explaining Collective Action Tendencies through Group-Based Anger and Group Efficacy." *Journal of Personality and Social Psychology* 87 (5): 649–64.