

Aspects of Blame

In which the nature of blame, blameworthiness, standing to blame and proportional blame are discussed

MARTA JOHANSSON WERKMÄSTER

DEPARTMENT OF PHILOSOPHY | LUND UNIVERSITY



Aspects of Blame

In which the nature of blame, blameworthiness,
standing to blame and proportional blame are discussed

Aspects of Blame

In which the nature of blame, blameworthiness,
standing to blame and proportional blame are discussed

Marta Johansson Werkmäster



DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the
Faculty of Humanities at Lund University to be publicly defended on 16th
of December 2023 at 11.00 in LUX:C126, Helgonavägen 3, Lund.

Faculty opponent
Matt King

Organization: LUND UNIVERSITY

Document name: Aspects of Blame

Date of issue: 2023-12-16

Author: Marta Johansson Werkmäster

Sponsoring organization:

Title and subtitle: Aspects of Blame: In which the nature of blame, blameworthiness, standing to blame and proportional blame are discussed

Abstract:

This thesis discusses what blame is, what it is for an agent to be blameworthy for having performed a certain act or omission, what makes an agent blameworthy for having performed a certain act or omission, whether the idea of standing to blame generalises to blame understood as something privately held, and what it is for blame to be proportionate. It provides original answers to these questions that move the current discussion of blame forward.

I argue for the novel claim that blame is a type of sentiment. When we blame someone, we are prone to experience various emotions, desires and thoughts in different circumstances with respect to her. Adoption of the blame sentiment does not involve interacting with anyone: it is "private". When an agent acts on the desires etc. that her blame prompts, the blame is "overt".

I develop and test a buck-passing account of blameworthiness. According to it, what it is for an agent to be blameworthy for having performed a certain act or omission is for there to be sufficient reason for anyone to blame her for having done that. I argue that what makes it the case that anyone has sufficient reason to blame her for having performed a certain act or omission is that she freely performed the act or omission despite being aware of the reasons which, taken together, made her act or omission one that it was wrong of her to perform or omit. In addition, I argue that we should understand "blame" in the analysis as referring to just private blame.

Many blame scholars believe that the idea of standing to blame applies to both private and overt blame. I argue against this claim by showing that common features included in the idea of the standing to blame do not apply to private blame in a way that would justify the claim that there are norms of standing for such blame. Similarly, I argue that because private and overt blame have a different characteristics, they are associated with different proportionality principles.

Key words: blame, sentiment, Strawson, reactive attitudes, fitting attitude analysis of value, buck-passing account of value, blameworthiness, volitional control, rational control, capacitarianism, emotions, degree, proportionality, standing to blame, normative power, privilege right, right kind of reason, wrong kind of reason, praise, praiseworthiness, rightness, wrongness.

ISBN (print): 978-91-89415-90-4

Language English

ISBN (digital): 978-91-89415-91-1

Number of pages: 202

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature



Date 2022-10-30

Aspects of Blame

In which the nature of blame, blameworthiness,
standing to blame and proportional blame are discussed

Marta Johansson Werkmäster



LUND
UNIVERSITY

Copyright Marta Johansson Werkmäster

The Joint Faculties of Humanities and Theology
Department of Philosophy

ISBN (print) 978-91-89415-90-4

ISBN (digital) 978-91-89415-91-1

Printed in Sweden by Media-Tryck, Lund University
Lund 2023



Media-Tryck is a Nordic Swan Ecolabel
certified provider of printed material.
Read more about our environmental
work at www.mediatryck.lu.se

MADE IN SWEDEN 

To Elliot

Table of Contents

Acknowledgment.....	11
List of Terms and Named Theses	13
Chapter 1 “Introduction”	13
Chapter 2 “On the Nature of Blame”	14
Chapter 3 “On What Makes Agents Blameworthy”	15
Chapter 4 “On Standing to Blame”	16
Chapter 5 “On Proportional Blame”	17
Chapter 6 “Testing the FA Analysis of Blameworthiness”	18
1 Introduction	19
1.1 On Strawson.....	21
1.2 The Fitting Attitude Analysis of Value	26
1.2.1 FA Blameworthiness	32
1.2.2 The WKR Problem.....	33
1.3 Methodology and Limitations.....	38
1.3.1 Methodology	38
1.3.2 Some General Limitations.....	39
1.4 Looking Ahead	40
2 On the Nature of Blame.....	43
2.1 Clarifying the Aim and Desiderata	44
2.2 Moral Blame, Non-Moral Blame, Epistemic Blame, Etc.	45
2.3 The Sceptical Claim.....	46
2.4 Blame Is Not Solely a Judgment, an Action or Being Angry	48
2.4.1 The Judgment Account.....	48
2.4.2 The Act Account.....	50
2.4.3 The Anger Account.....	52
2.4.4 Summary	58

2.5	Hybrid Accounts.....	58
2.5.1	Sher's Account	58
2.5.2	Scanlon's Account.....	60
2.5.3	Summary	65
2.6	The Functionalist Account.....	66
2.7	The Process Account.....	69
2.8	The Sentiment Account	72
2.8.1	Sentiments and Blame.....	72
2.8.2	How the Sentiment Account Avoids Previous Objections .	78
2.8.3	The Sentiment Account and Badness	81
2.9	Conclusion	83
3	On What Makes Agents Blameworthy	85
3.1	Attributability, Accountability, Etc.	87
3.2	<i>De Dicto</i> and <i>De Re</i>	89
3.3	Two Cases	90
3.4	The Witting Wrongdoing View.....	94
3.5	The Wrong-Making View.....	101
3.6	Challenges	110
3.6.1	Blameworthiness and Wrongdoing	110
3.6.2	The Wrong-Making View** and Excuses	114
3.6.3	Degrees of Blameworthiness.....	117
3.7	Conclusion	120
4	On Standing to Blame	121
4.1	Private Blame and Overt Blame	123
4.2	Core Features of Standing.....	125
4.3	Scepticism About the Idea of Standing to Blame	
	Someone Privately	129
4.3.1	The Nature of Standing and Private Blame	129
4.3.2	Conditions of Standing and Private Blame.....	130
4.3.3	Hypocritical Private Blame and Wrongness.....	131
4.3.4	Guidance and Private Blame	134
4.3.5	Summary	137
4.4	Conclusion	138

5	On Proportional Blame.....	139
5.1	Private Blame and Overt Blame Again	140
5.2	Proportional Private Blame and Proportional Overt Blame	146
5.2.1	Proportional Private Blame	147
5.2.2	Proportional Overt Blame.....	148
5.3	Some Practical Implications	154
5.4	Conclusion	155
6	Testing the FA Analysis of Blameworthiness	157
6.1	FA Blameworthiness and the Ethics of Blame	158
6.1.1	FA Blameworthiness and Standing to Blame.....	158
6.1.2	FA Blameworthiness and Proportional Blame	159
6.1.3	Summary and Generalising the Challenge.....	160
6.2	The WKR Problem Again.....	161
6.2.1	The RKR and the WKR	162
6.2.2	Dissolving the Tension	164
6.3	Conclusion	167
7	Summary and Critical Discussion	169
8	Appendix: On Praise	175
8.1	On the Nature of Praise	175
8.2	On What Makes Agents Praiseworthy.....	178
8.3	On Standing to Praise.....	184
8.4	On Proportional Praise	187
8.5	On Fitting Praise	190
8.6	Summary	191
9	References	193

Acknowledgment

I welcome this opportunity to express my gratitude to some of the people who have given me help in writing this thesis or in making writing it possible.

Those with whom I have had the good fortune to interact, and who have had a direct impact on what I have to say in this thesis, include the following: David Alm, Henrik Andersson, Agnès Sophia Constance Baehni, Monika Betzler, Gunnar Björnsson, Olle Blomberg, Francesca Bunkenborg, Julien A. Deonna, Leo Eisenbach, Anton Emilsson, Mohammadhadi Fazeli, Andrés Garcia, Mattias Gunnemyr, Fritz Gåvertsson, Jessica Isserow, Ingvar Johansson, Jens Johansson, Yulia Kanygina, Benjamin Kiesewetter, Jiwon Kim, Marianna Leventi, Kasper Lippert-Rasmussen, Leonhard Menges, Per Millan, Shervin MirzaeiGhazi, Gloria Mähringer, Francesco Orsi, August Olsen, Robert Pál Wallin, Björn Petersson, Wlodek Rabinowicz, Paul Robinson, Paul Russell, Toni Rønnow-Rasmussen, Signe Savén, Thomas Schmidt, David Shoemaker, Justin Snedegar, András Szigeti, Matthew Talbert, Daniel Telech, Fabrice Teroni, Patrick Todd, Caroline Torpe Touborg, Alexander Velchikov, Jakob Vikner, Shawn Wang and Jakob Werkmäster.

Of the people just mentioned, nine deserve special recognition.

David Alm, András Szigeti, and Matthew Talbert were my supervisors for this research. I am very grateful for their feedback on earlier drafts of the thesis. Answering their comments improved the thesis significantly, and I have learned a lot along the way. Special thanks are owed to my main supervisor, András Szigeti, whose constant interest and encouragement made this thesis possible.

Leonhard Menges was my mock-defence opponent. I am very grateful for his comments. Answering them also improved the thesis significantly.

Apart from having been my office buddy for five years, Mattias Gunnemyr read parts of the draft for the mock-defence and provided written feedback on it that I appreciated a lot.

Over the course of my time as a PhD student at Lund University, Jakob Werkmäster has, in a supportive manner, helped me in my efforts to develop and refine my ideas about blame. In addition, he has co-authored several papers on (among other things) blame with me. I am very grateful for that support and those

collaborations. Not only that, but our personal relationship means the world to me – I have never experienced an intense love like ours before (<https://open.spotify.com/track/77zz8fPUkrXjqK5sFyoHhm?si=6a56b6201b0f4a92>).

I was part of the Lund Gothenburg Responsibility Project (LGRP), of which Paul Russel is the PI. I want to thank him, and the group's members, for organising reading groups, workshops and conferences that influenced many of the ideas presented in this thesis.

Paul Robinson proofread the thesis. I am very grateful for his careful editing.

You do not become a PhD student at the Department of Philosophy, Lund University, nor intend to explore emotive theories of blame and the fitting attitude analysis of blameworthiness in your thesis out of the blue. Thanks must go to my former supervisors at the Department of Philosophy, Uppsala University, Jens Johansson and Andrew Reisner, for encouraging me to apply for the PhD position I currently hold, and (especially Andrew) for introducing me to the fitting attitude analysis of value. Also, thank you, Rafaelle Rodogno (Aarhus University) for introducing me to the philosophy of emotions during my time as an Erasmus student at the University of Aarhus, and for writing me a much appreciated reference letter.

I want to express my gratitude to Ingela Byström, Anna Cagan Enhörning, Petter Johansson, Martin Jönsson, Tomas Persson, Eva Sjöstrand, Annah Smedberg-Eivers, Erika Sombeck, Tobias Hansson Wahlberg, Annika Wallin and Anna Östberg for being such nice and helpful colleagues. In addition, thank you, Dan Egonsson and Cathrine R. Felix for making me feel at home in Lund and in the Philosophy Department. Lastly, thank you my friends, my family and my cat Ville for your support and love throughout the years.

My main debt of gratitude is to my mother, Malgorzata Wojnar-Johansson, and my father, Nils Johan Johansson. Thank you for helping me to move to different places; for encouraging me to finish my tasks and pursue my interests; and for the inspiration you have provided throughout the years.

I want to end this section by thanking the person to whom this dissertation is dedicated, my son Elliot Johansson Werkmäster.

List of Terms and Named Theses

In this section, I present a list of terms and named theses relevant for each chapter in this thesis.

Chapter 1 “Introduction”

Conclusive reason: there is conclusive reason for an agent A to ϕ if, and only if, A 's *pro tanto* reasons to ϕ are stronger than A 's *pro tanto* reasons to not- ϕ .

The FA analysis (generic): an object x is (dis)valuable if, and only if, it is fitting to (dis)favour x .

The FA analysis: an object x is (dis)valuable to degree n if, and only if, (and because) it is fitting for anyone to (dis)favour x to degree n .

The FA analysis (RKR): An object x is (dis)valuable to degree n if, and only if, (and because) there is a sufficient RKR for anyone to (dis)favour x to degree n .

FA blameworthiness: an agent A is blameworthy to degree n for having ϕ -ed if, and only if, (and because) it is fitting for anyone to blame A for having ϕ -ed to degree n .

Normative reason: a *normative reason* for an agent A to ϕ is a distinct fact, or set of facts, that counts in favour of A ϕ -ing.

Pro tanto reason: a *pro tanto* reason for an agent A to ϕ is a normative reason for A to ϕ with strength S .

Sufficient reason: there is sufficient reason for an agent A to ϕ if, and only if, the *pro tanto* reasons for A to ϕ are not outweighed by A 's *pro tanto* reasons to not- ϕ .

Wrong kind of reason problem (WKR problem): the WKR problem is that the FA analysis is extensionally inadequate – e.g., in some cases an agent has a sufficient reason to direct a pro-attitude upon an object that clearly lacks value.

Chapter 2 “On the Nature of Blame”

The act account of blame: an agent A blames another agent B for having ϕ -ed if, and only if, A boos, or punishes, B for having ϕ -ed.

The anger account of blame:

Other-blame: an agent A blames another agent B for having ϕ -ed, and where it is not the case that $A=B$, if, and only if, A is angry with B for having ϕ -ed.

Self-blame: an agent A blames herself for having ϕ -ed if, and only if, A feels guilt over the fact that she has ϕ -ed.

The functionalist account about the nature of blame: normally, anything that has the function of, say, signalling one’s commitment to a certain norm N is blame.

The judgment account about the nature of blame: an agent A blames another agent B for having ϕ -ed if, and only if, A judges that B has a stain in her moral record (about her lifetime moral worth) in virtue of having ϕ -ed.

Overt blame: to intentionally act on the desires, emotions, or thoughts the sentiment of blame prompts.

Private blame: to merely harbour the sentiment of blame.

The process account of the nature of blame: an agent A blames another agent B for having ϕ -ed if, and only if, A undergoes a certain process P consisting of thoughts, emotions, motivations and actions, and experience it as a process.

Scanlon’s account of the nature of blame: an agent A blames another agent B for having ϕ -ed if, and only if, (i) A judges that B has impaired their relationship by having ϕ -ed, and (ii) A modifies her relationship to B in an appropriate way to the judgment.

The sentiment account of the nature of blame:

Self-blame: an agent A blames another agent B for having ϕ -ed, and where $A=B$, if, and only if, A has acquired a specific disposition β_1 .

Other-blame: an agent A blames another agent B for having ϕ -ed, and where it is not the case that $A=B$, if, and only if, A has acquired a specific disposition β_2 .

Sher's account of the nature of blame: an agent A blames another agent B for having ϕ -ed if, and only if, (i) A believes that B has acted wrongly, or badly, by ϕ -ing, and (ii) desires that B had not ϕ -ed.

Chapter 3 “On What Makes Agents Blameworthy”

Indirect blameworthiness: an agent A is indirectly blameworthy for having performed a particular action or omission ϕ if, and only if, A is blameworthy for ϕ by way of being blameworthy for some other action or omission ψ , of which ϕ is or was a foreseeable consequence of ψ .

Objectivism about rightness and wrongness: rightness and wrongness depend on *all* the facts and not only facts epistemically available to agents.

Rational control: the sort of control we exercise by being receptive and reactive to certain reasons, by forming, revising (etc.) our attitudes in light of our awareness of certain facts.

Standard view of excuses: An agent A is excused for performing a certain action or omission ϕ if, and only if, (i) it was wrong for A to ϕ and (ii) A is not at all blameworthy for having ϕ -ed.

Subjectivism about rightness and wrongness: rightness and wrongness depend on facts epistemically available to agents.

Volitional control: the sort of control we exercise when we ϕ deliberately, or ϕ for more or less whatever reason we think is sufficient (e.g., to win a bet or to impress someone) and decide when to ϕ (e.g., now or later). Whether to ϕ is “up to the agent”.

The witting wrongdoing view: an agent A is blameworthy for having performed a particular action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) A ϕ -ed despite consciously believing it to be wrong to ϕ .

The wrong-making view: an agent A is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) A 's motivating reason(s) for ϕ -ing coincides with the normative reason(s) that makes it wrong for A to ϕ .

*The wrong-making view**: an agent A is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) A 's motivating reason(s) for ϕ -ing coincides with the normative reason(s) that makes it wrong for A to ϕ or A 's ϕ -ing is rationalised by A 's moral indifference.

*The wrong-making view***: an agent A is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) A ϕ -ed despite, at the time of action or omission, being aware of the fact(s) that constitutes a normative reason(s) which, taken together, makes it wrong for A to ϕ .

*The wrong-making view****: an agent A is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) A did ϕ despite, at the time of action or omission, being aware of the fact(s) that constitutes a normative reason(s) which, taken together, makes it wrong for her to ϕ , and (iii) A could and should have believed that ϕ -ing is wrong.

Chapter 4 “On Standing to Blame”

The avoidance of meddling condition: an agent, B , points to facts showing that another agent, A , is an outsider to the underlying matter in which A intervenes. Being an outsider to the underlying matter is what undermines A 's standing to blame B .

The commitment view: A lack of commitment to a certain norm N undermines an agent's standing to blame other agents for violations of N .

The absence of hypocrisy condition: see the moral equality view and the commitment view.

The moral equality view: in virtue of rejecting the principle concerning the equality of persons, that morality applies equally to us all, with respect to a certain norm N , undermines an agent's standing to blame other agents for violations of N .

Power-right: normative powers change agents' normative landscape by for instance creating, intensifying or undercutting, reasons. One needs to have a certain authority to successfully exercise normative powers. The idea is that by blaming agent B (granting that agent A has the proper authority), A gives B a *pro tanto*

reason to apologise, explain her behaviour, intensify B 's *pro tanto* reason to apologise, and such alike.

Privilege-right: an agent A has a privilege-right to ϕ if, and only if, A has no obligation not to ϕ .

Chapter 5 “On Proportional Blame”

*Disproportionate overt blame conveyed**: an act of overtly blaming a blameworthy agent B for having ϕ -ed to degree_{conveyed} n is disproportionate_{conveyed} if, and only if, n exceeds, or is less than, the intensity i of private blame that it is fitting to direct upon B for having ϕ -ed.

*Disproportionate overt blame harm**: an act of overtly blaming a blameworthy agent B for having ϕ -ed to degree_{harm} n is disproportionate_{harm} if, and only if, n exceeds the goodness brought about by the overt blame in terms of B taking responsibility for her having ϕ -ed.

Disproportionate private blame: an instance of privately blaming a blameworthy agent B for having ϕ -ed with intensity i is disproportionate if, and only if, i exceeds, or is less than, the intensity i of private blame that it is fitting to direct upon B for having ϕ -ed.

*Proportionate overt blame conveyed**: an act of overtly blaming a blameworthy agent B for having ϕ -ed to degree_{conveyed} n is proportionate_{conveyed} if, and only if, n matches the intensity i of private blame that it is fitting to direct upon B for having ϕ -ed.

*Proportionate overt blame harm**: an act of overtly blaming a blameworthy agent B for having ϕ -ed to degree_{harm} n is proportionate_{harm} if, and only if, n does not exceed the goodness brought about by the overt blame in terms of B taking responsibility for her having ϕ -ed.

Proportionate private blame: an instance of privately blaming a blameworthy agent B for having ϕ -ed with intensity i is proportionate if, and only if, i matches the intensity i of private blame that it is fitting to direct upon B for having ϕ -ed.

Chapter 6 “Testing the FA Analysis of Blameworthiness”

Action-oriented values: values that correspond directly to *just* an action.

Attitude-oriented values: values that correspond directly to just an attitude or also to an attitude.

*FA blameworthiness**: an agent A is blameworthy to degree n for having ϕ -ed if, and only if, (and because) it is fitting for anyone to privately blame A for having ϕ -ed to degree n .

1 Introduction

You are reading the news. The front page says that a black American man has been killed by police during an arrest.

It began with a report, to the police, saying that a man had used counterfeit bills. In response, the police set out to arrest the man. During the arrest, the man, according to one of the police officers, resisted being placed in the car. One of the police officers therefore forced the man to the ground and put a knee on his neck – this, despite the man having been handcuffed.

The police officer kept his knee on the neck for about nine minutes, even though the man repeatedly said “I can’t breathe”. Some bystanders, who also filmed the event, pleaded with the police officer to remove his knee, because they saw that he was about to kill the man. The officer did not respond by lifting his knee. When the medics finally came, he lifted it, leaving behind a dead man.

This is a rough account of how George Floyd’s life ended on May 25, 2020. The story is not unique. It resembles the way many other black Americans in the United States have died due to police brutality: to name a few, Tanisha Anderson, Michael Brown, Tamir Rice, Gabriella Nevarez, Akai Gurley, Michelle Cusseaux, Eric Garner, Janisha Fonville, Freddie Gray, Philando Castine, Botham Jean, Stephon Clark, Aura Rosser, Atatiana Jefferson, Breonna Taylor, Manuel Ellis, Andre Hill, and Daunte Wright.

The case of George Floyd is not only tragic and upsetting. It also raises interesting and not yet adequately explored philosophical questions about blame.

Suppose some people are angry with the police officer in response to his conduct. Are these angry responses blame? Or is blame something calmer or more moderate? According to some accounts of the nature of blame, to blame someone is (more or less) to be angry with her. But for the defenders of other accounts, blame is not always something angry.

Further, suppose the police officer, in response to being accused of police brutality by protesters, claims that he was *just* deliberately doing what he believed was right.¹ And assume he is speaking truthfully. Does this reply show that he is not blameworthy for killing George Floyd? Intuitively, the answer is “No”. However, on one popular account on what makes agents blameworthy, the answer can be “Yes”.

Furthermore, suppose there is another police officer – one charged with committing the same kind of police brutality at an earlier time who has not made adequate amends since then. And suppose he blames the police officer in our earlier story. Does the second officer lack the standing, or right, to blame his colleague? According to most contributors to the debate, the answer would probably be “Yes”. It is less clear, however, whether these contributors think that the second police officer lacks the standing to overtly blame (some action) the first police officer or to privately blame (some attitude) him, or both. Further, it is unclear what is meant by “standing” or “right”.

Finally, suppose that many people have blamed the police officer, and on repeated occasions – say, thousands or even millions of people have sent him tweets expressing their blame – without him having made adequate amends at any point in response. Does this fact suggest that it is not fitting for you to blame the police officer because your blame would be “too much”? Can blame be too much? If yes, how and why? There are no clear answers to these questions in the philosophical literature on blame.

This thesis will take up issues such as those sketched. More precisely, it aims to answer the following five questions:

1. What is the nature of blame?
2. What makes an agent blameworthy for having performed a particular action or omission?
3. Do norms of standing to blame apply to private blame?
4. What is it for an instance of blame to be proportionate?

¹ Here I mean overall morally right, not *prima facie* right.

5. How do the ideas of standing to blame and proportional blame relate to the fitting attitude analysis of blameworthiness? For example, are they in tension?

My purpose in this introductory chapter is to set out the most important assumptions and background information on which my arguments are based. In addition, I would like to outline the structure of the dissertation. Along the way, I will give more substance to research questions (1)-(5) above.

1.1 On Strawson

Peter F. Strawson's seminal work "Freedom and Resentment" (1962) is the point of departure for many contemporary theorists of blame. This will be clear in the chapters to come. It is therefore helpful to provide a brief overview of the paper's basic ideas as a backdrop to the chapters to come.

To narrow things down, I will focus on the way in which some of Strawson's ideas have been used by various blame scholars when answering questions (1) and (2) presented in the previous section. Few, if any, blame theorists have alluded in any significant way to Strawson's work when answering questions (3)-(5), and therefore, I will remain silent on those questions in this section.

I aim to remain as neutral as possible with respect to various interpretations of Strawson and to work with "standard" interpretations of the main ideas in Strawson's essay. By "standard interpretations", I mean those interpretations that are most commonly made in the literature.

Before describing how some of Strawson's ideas have been used in the debate about blame, let me provide a brief overview of the essay.

In "Freedom and Resentment", Strawson aims to settle the dispute between *pessimists* and *optimists*. By "pessimists" Strawson means those who accept that the truth of determinism would imply that the concept of blame has no application and that the practice of blame is always unjustified.² By "optimists", on the other hand, Strawson means those who believe that the truth of determinism would imply neither that the concept of blame has no application nor that the practice

² In short, determinism is the thesis that the past and the laws of nature determine a unique future. For a good exposition of determinism, see van Inwagen (1975).

is always unjustified. Strawson's optimist typically invokes consequentialist considerations. Strawson states that the optimist

[...] undertakes to show that the truth of determinism would not shake the foundations of the concept of moral responsibility and of the practices of moral condemnation and punishment, he typically refers, in a more or less elaborated way, to the efficacy of these practices in regulating behaviour in socially desirable ways. (Strawson 1962: 22)

Between these opposing positions Strawson presents and discusses a further position, namely that of the *moral sceptic*. By "moral sceptics" Strawson means those who hold that the concept of blame is inherently confused, and that blame is unjustified, regardless of whether the thesis of determinism is true.

In the essay, Strawson argues that we can settle the dispute between pessimists and optimists (and moral sceptics) by attending to "what it is actually like to be involved in ordinary inter-personal relationships, ranging from the most intimate to the most casual" (Strawson 1962: 7). When we attend to our actual practice of holding one another responsible, it becomes evident, according to him, in what way the thesis of determinism is relevant to that practice. He concludes that the thesis of determinism plays no role in our everyday practices of holding one another responsible.

I now move on to describe how Strawson's ideas have been used by some blame scholars to answer questions (1) and (2). Before I do that, however, I want to stress that Strawson did not aim to provide a comprehensive theory of, say, the nature of blame. He rarely uses the term "blame" in his essay. In addition, he even confesses to being aware that his language and ideas, because they are commonplaces, are "quite unscientific and imprecise" (Strawson 1962: 5). What Strawson intended to do – and what he succeeded in doing, I believe – was rather to shift the focus from metaphysical questions about free will to the attitudes with which we blame and hold each other responsible.

Strawson famously distinguished between *participant attitudes* and *objective attitudes*. By the former he means

the non-detached attitudes and reactions of people directly involved in transactions with each other; of the attitudes and reactions of offended parties and beneficiaries;

of such things as gratitude resentment, forgiveness, love, and hurt feelings. (Strawson 1962: 5)

In other words, his focus is on the attitudes we have towards people we view as agents or our peers. The reactive attitudes subdivide, according to Strawson, into three groups:

Personal reactive attitudes: attitudes that are directed toward others' treatment of oneself, with that treatment triggering reactive attitudes such as resentment and gratitude.

Vicarious reactive attitudes: attitudes that are directed toward others' treatment of others', with that treatment triggering reactive attitudes such as indignation and approbation.

Self-reactive attitudes: attitudes that are directed toward one's own treatment of others, with that treatment triggering reactive attitudes such as guilt and shame.

Strawson does not say much more about the nature of the reactive attitudes. In the blame literature, it is commonly assumed that reactive attitudes are emotions.³ However, this characterisation is not very informative in itself, because there are dramatically divergent views on the nature of emotions, and each view gives a different verdict on how the nature of the reactive attitudes (assuming they are emotions) ought to be understood. While a judgmentalist about emotions would hold that emotions such as resentment are more or less equivalent to certain judgments, some non-judgmentalists would argue that reactive attitudes are feelings (possibly constituted by, or even reducible to, bodily changes) or relevantly similar to perceptions.⁴ In addition, it is unclear whether "emotion" here refers to something occurrent, an episode, or something dispositional, or to a process, or a mood – or to some combination of these.

In the blame literature, it is common to account for blame in terms of the reactive attitudes, or even to reduce blame to reactive attitudes. For example, R. Jay

³ See, for instance, Menges (2017) and Shoemaker (2015).

⁴ For some judgmentalist views of emotions, see Nussbaum (2001) and Solomon (1976). For some non-judgmentalist views of emotions, see James (1884) and Tappolet (2016).

Wallace (1994) states: “I propose that blame involves a susceptibility to the reactive emotions” (Wallace 1994: 12), and Susan Wolf (2011) similarly claims:

More precisely, my ordinary use of the term associates blame with a certain kind of negative emotional attitude toward the object of blame – resentment, indignation, and guilt, as well as righteous anger, fall within the family of these attitudes, but the mere (or not so mere) absence or withdrawal of good will does not. (Wolf, in Wallace et. al 2011: 335)

In Chapter 2 (§2.2.3) of this thesis, I will argue against this Strawson-inspired view of the nature of blame. My main objection to it is that, at least as construed by Wallace and Wolf at any rate, it is too narrow. We tend to feel emotions other than mere anger and guilt when we blame others or ourselves. However, at the end of the chapter (§2.8), I will propose a view of the nature of blame that is nevertheless “Strawsonian” insofar as it is affective and involves susceptibility to some reactive attitudes, such as resentment, indignation and guilt.⁵

By “objective attitude” Strawson means the detached attitudes that we adopt towards people when we see them as objects of social policy or treatment – i.e., as not our peers. The people Strawson seems to have in mind here are those suffering from severe mental illness, children, and the “psychologically abnormal”. He summarises nicely what he means by “objective attitude” here:

To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided, though this gerundive is not peculiar to cases of objectivity of attitude. The objective attitude may be emotionally toned in many ways, but not in all ways: it may include repulsion or fear, it may include pity or even love, though not all kinds of love. (Strawson 1962: 9-10)

Strawson does not provide a clear threshold delineating when an individual is an object of social policy and when she is not. For example, when does a child go

⁵ Others have linked the nature of blame to alternative features mentioned in Strawson’s essay. For example, Pamela Hieronymi (2004) bases her account of blame on Strawson’s idea that we react to people’s quality of will when we blame them, and that we additionally demand a certain degree of goodwill from them. In a moment, I will say more about quality of will. Briefly, her view is that to blame is to judge that a person has manifested insufficient goodwill.

from being an object of social policy to being our peer or a fellow agent? Regardless of this and similar concerns, the crucial and, I believe, intuitive point is that we do not tend to react in the same way towards those we view as agents or our peers, on the one hand, and those we do not view as agents or our peers, on the other.

That said, Strawson has repeatedly argued that although we naturally tend to react with participant attitudes toward people we view as agents and with objective attitudes toward people we view as non-agents, the objective attitude

[...] is also something which is available as a resource in other cases too. We have this resource and can sometimes use it; as a refuge, say, from the strains of involvement; or as an aid to policy; or simply out of intellectual curiosity. (Strawson 1962: 10)

He adds that although we can deliberately adopt the objective attitude toward agents, and use it as a “resource” from time to time, we can only do so temporarily. He explains:

I do not think it is a point of view or position which we can hold, or rest in, for very long. The price of doing so would be higher than we are willing, or able, to pay; it would be the loss of all human involvement in personal relationships, of all fully participant social engagement. (Strawson 1985: 27-8)

Moving to another core idea of Strawson’s, consider the following, now famous, case that he provides:

If someone treads on my hand accidentally, while trying to help me, the pain may be no less acute than if he treads on it in contemptuous disregard of my existence or with malevolent wish to injure me. But I shall generally feel in the second case a kind and degree of resentment that I shall not feel in the first. (Strawson 1962: 6)

Strawson suggests that although both agents produce the same bad or harmful outcome – that is, stepping on my hand – that does not say much about whether or not we would react with the reactive attitudes towards them on the basis of that. Rather, what elicits the reactive attitudes, according to him, is whether they stepped on my hand with some bad quality of will or not. Not only that, but one also needs to consider whether, in stepping on my hand with that bad quality of will, they violated some demand, or requirement. According to Strawson, we

[...] demand some degree of goodwill or regard on the part of those who stand in these relationships to us, though the forms we require it to take vary widely in different connections. (Strawson 1962: 6-7)

Strawson connects his unique view of excuses to related ideas about quality of will. According to him, successful (i.e., fully effective) excuses show that what may have seemed like poor quality of will really was not. For example, it might appear that my stepping on your foot shows that I do not care about you. However, if I demonstrate to you that I stepped on it accidentally, and that I do indeed care about you, it will be false to say that poor quality of will was manifested by my action, and therefore, I will be fully excused for my action, according to Strawson.

It is not entirely clear what Strawson means by “quality of will”, “demand”, and the associated terms above. He does not provide more information about it in the essay, nor does he do so elsewhere. In Chapter 3, I will critically discuss some current and possible explications of “quality of will” (or rather quality of *mind*) and “demand”, and defend one of them.

To conclude this section, the main Strawsonian ideas we need to hold in our minds as background information in the chapters to come are the following: first, his notion of reactive attitude, and second, his idea that the agent’s quality of will, or mind, is relevant to blameworthiness.

1.2 The Fitting Attitude Analysis of Value

In this section, I present the theory many arguments in this thesis are based on. In addition, it is this theory I will consider when answering research question (5).

It is commonly believed that there is a relation between value properties and fitting pro- or con-attitudes or responses. For example, we think it is fitting to admire the admirable, fitting to love the loveable, fitting to protect the good, fitting to choose the choiceworthy, fitting to blame the blameworthy, and so on. Defenders of the *fitting attitude analysis of value* (FA analysis of value) can be interpreted as developing this widely held belief into a full-blown analysis of value. In its most generic form, the FA analysis of value can be formulated as follows:

The FA analysis (generic): An object x is (dis)valuable if, and only if, it is fitting to (dis)favour x .

The FA analysis of value, thus stated, clearly needs some unpacking. Below I briefly present what I mean by “(dis)valuable”, “fitting”, and “(dis)favour”.

“(Dis)valuable” refers to what is valuable *simpliciter*. The notion that an object is (dis)valuable *simpliciter* should be distinguished both from the idea that it is (dis)valuable-for someone (e.g., my child’s painting is valuable for me) and from the idea of *attributive-value* (e.g., bad knife). Further, (dis)value *simpliciter* can be either *final* (happiness is typically thought of as being finally good) or *instrumental* (if money is a means to happiness, it is instrumentally good). Furthermore, something can be (dis)valuable *simpliciter* either *in a respect* or *overall*. For example, a collection of essays which includes one essay that is very stimulating, even though the rest are dull, might be good in the respect that it contains one stimulating essay but bad overall. In this thesis, I will be concerned with (dis)valuable *simpliciter*, final (dis)value and the overall notion, if not stated otherwise.

I understand “fitting”, which is the normative part of the FA analysis of value, in terms of there being a *sufficient reason* to (dis)favour. There is sufficient reason for an agent A to ϕ if, and only if, the *pro tanto* reasons for A to ϕ are not outweighed by A ’s *pro tanto* reasons to not- ϕ .

It may be wondered why I have postulated that the FA analysis of value stipulates that there is a sufficient reason and not, for example, a *pro tanto* reason or a *conclusive* reason. A *pro tanto* reason for an agent A to ϕ is a normative reason for A to ϕ with strength S . And there is conclusive reason for an agent A to ϕ if, and only if, A ’s *pro tanto* reasons to ϕ are stronger than A ’s *pro tanto* reasons to not- ϕ . The answer, in short, is that claiming that an object is (dis)valuable when there is a *pro tanto* reason to (dis)favour it would generate too many values. We have *pro tanto* reasons to favour all sorts of things, and the reasons do not always show that an object is valuable. Analysing values in terms of conclusive reasons, in contrast, seems to generate too few values. We rarely have a conclusive reason to (dis)favour something.

While many advocates of the FA analysis of value understand the normative part in terms of reasons, not all do.⁶ It is popular nowadays to take “fitting” to be basic and to understand the normative element just to mean “fitting”.⁷ I will not evaluate this explication in this thesis. Nor will I defend the reason-version I deploy, or *the buck-passing account* I deploy. My aim is more modest: I want to formulate an informative FA analysis of blameworthiness and test it by appealing to some central claims made in the ethics of blaming.⁸ The ethics of blame lays down a series of conditions relevant to whether or not an instance of blame is permissible.⁹ Everyone seems to agree that two such conditions are that blame is *proportional* and that the blaming agent has in fact *standing* to blame the blameworthy agent. If your blame is not proportional, or if you lack standing to blame the blameworthy agent, then that speaks against the permissibility of that blame. In Chapter 6, I show that the FA analysis of blameworthiness and central claims made in the ethics of blaming are in tension depending on how we specify the FA analysis of blameworthiness and the central claims made in the ethics of blaming.¹⁰

I want to make it clear that just because I will formulate an FA analysis of blameworthiness and test it, this should not be taken to mean that I accept it as a background assumption or that I believe that the FA analysis of value more generally speaking would be true (in a moment, I will present some severe objections to the FA analysis of value). Rather, I am interested in testing and developing the FA analysis of blameworthiness because it forms part of an influential wider analysis of value. Further, I believe it is interesting to investigate whether certain specifications of the analysis are plausible. If they are not, then that can be seen as speaking against the FA analysis of value *pro tanto*, in the sense

⁶ For defenders of the reason interpretation of the FA analysis of value, or the buck-passing account of value, see in particular Scanlon (1998), Parfit (2011), Rabinowicz and Rønnow-Rasmussen (2004), Skorupski (2010a), Schroeder (2010), and Cosker-Rowland (2019).

⁷ See especially McHugh and Way (2016; 2022).

⁸ I am not the first who has formulated an FA analysis of blameworthiness – see, for example, Skorupski (2010a) and King (2012). One can say that, in this thesis, I stand on these scholars’ shoulders and ask, given that we can formulate an FA analysis of blameworthiness, how should we fill in the details of it and how should we relate it to the debate on the ethics of blaming. Below, I say more about the details I aim to fill in this thesis.

⁹ For more on the ethics of blame, see Coates (2020) and Scanlon (2008).

¹⁰ Below, I will present the FA analysis of blameworthiness.

that we cannot rely on it to make sense of some value properties, such as “blameworthy”. Finally, as I will show in Chapter 6, I believe that the novel problems I will raise against the FA analysis of blameworthiness can be generalised to FA analyses of other value properties. Put differently, the results in this thesis should be of interest to all scholars who believe that there is a relation between value properties and fitting attitudes or responses.

Advocates of the FA analysis of value rarely explicate the attitudinal part of the analysis.¹¹ Although it is called the fitting *attitude* analysis of value, the attitudinal part has referred not just to attitudes, but also to actions, or some combination of attitudes and actions. For example, an object *x* is good if, and only if, it is fitting to protect, preserve, like or/and favour *x*. The only guideline as regards what the attitudinal part of the FA analysis of value should refer to is that the attitude, or action, should in some direct way correspond to the value property we are concerned with – e.g., the attitude “admiration” directly corresponds to the value property “admirable”, the action “use” directly corresponds to the value property “usable”, and the responses/attitudes “promote”, “protect”, “desire”, “recommend” and “favour” directly correspond to the value property “good”.

In Chapter 6, I will discuss how we should understand the attitudinal part of the FA analysis of value, in particular the attitudinal part of the FA analysis of blameworthiness. I will provide some reasons for interpreting it just in terms of fitting attitudes.¹²

The FA analysis (generic) can be improved. As stated, it is silent on questions about *for whom* it is fitting to (dis)favour, and about the order of explanation, and about degrees of value.

¹¹ Except, perhaps, (*neo*)*sentimentalists*. They typically argue that the attitudinal part denotes attitudes, or rather, emotions. D’Arms and Jacobson (2000b) seem to endorse such a view. Consider what they write here: “[...] an important set of evaluative concepts (or terms or properties) is best understood as invoking a normative assessment of the appropriateness (or merit or rationality) of some associated emotional response” (D’Arms & Jacobson 2000b: 729).

¹² Put differently, I will argue against FA analysts who think that the attitudinal part can refer to actions, attitudes, or some combination of actions and attitudes and not FA analysts or (*neo*)*sentimentalists* who think that the attitudinal part should strictly refer to fitting attitudes or emotions (see fn 11). Scanlon (1998), Parfit (2011), Cosker-Rowland (2019), Berker (2022) to name a few FA or BPA scholars, seem think that the response component can refer to actions, attitudes, or some combination of both.

Although, usually, it is not explicitly stated *for whom* it is fitting to (dis)favour, it is commonly assumed that it is fitting for no one in particular, i.e., that it is fitting for anyone or everyone to do the (dis)favouring.¹³ Related to this, there are those who claim that it need only be fitting for agents who have satisfied a certain *epistemic familiarity condition* to (dis)favour.¹⁴ The idea is that it need not be fitting for agents who have never contemplated a particular object to (dis)favour it. In this thesis, I will assume that the notion we are working with is that it is fitting for anyone to do the (dis)favouring. I will not take a stance on whether, in the analysis, agents need to satisfy a certain epistemic familiarity condition. Nothing important hinges on whether we include such a condition or do not do so.

While some FA theorists take the object's (dis)value to be metaphysically prior, others shift the order of explanation and take the fact that it is fitting to (dis)favour the object to explain why the object is (dis)valuable. I would like to make it transparently clear that I will assume that the right-hand side of the bi-conditional is metaphysically prior to the left-hand side. However, once again, nothing important hinges on this. I believe that much of what I say in this thesis will apply to those who assume the reverse order of explanation is correct.

There is a debate among FA analysts over whether we should make sense of value's degree of strength by appeal to the normative part of the FA analysis of value – i.e., in terms of “fittingness” – or the attitudinal part – i.e., in terms of “(dis)favour”. More precisely, this issue concerns whether an object's (dis)value is a matter of *how fitting* it is to (dis)value it or a matter of *how much (dis)favouring* it is fitting to direct towards it.¹⁵

I will not seek to offer a knock-down argument explaining why we should prefer the view that a value's degree of strength is a matter of how much (dis)favouring it is fitting to direct towards the object. Instead, I will merely provide one compelling reason why making sense of values' degree of strength by appealing to

¹³ See, for example, McHugh and Way (2016), Orsi (2015), Rabinowicz (2013), Cosker-Rowland (2019), and Schroeder (2010).

¹⁴ See, for example, Bykvist, (2009), Chisholm (1986), Lemos (2011), and Zimmerman (2011b).

¹⁵ Of course, it could be argued that we should make sense of values' degree of strength by appealing to the normative *and* that attitudinal part of the FA analysis of value (e.g., Coates (2019)). This view is not popular and that is why I will not discuss it here.

the normative part of the FA analysis of value seems to be a non-starter.¹⁶ For, in this thesis (especially in Chapters 3 and 6), I will assume that we make sense of an agent's degree of blameworthiness by appealing to how much blame it is fitting to direct upon her for having ϕ -ed.

Above, I wrote that I take “fitting” to mean “sufficient reason”. Sufficient reasons are usually understood not to admit of degrees. A reason is either sufficient or it is not – the question is binary and there is no third position in between. This suggests it would be unwise to try to account for an object's degree of value by appeal to how fitting it is to favour it.

One might argue that making sense of values' degrees of strength by appealing to the attitudinal element of the FA analysis of value is not very plausible either. Before presenting a challenge to what (we might call) the attitudinal account, let me unpack it a bit.

I have said that the attitudinal part of the FA analysis of value has referred to attitudes, actions, or some combination of actions and attitudes. We typically make sense of attitudes' degree of strength by noting their intensity – i.e., the more intense an attitude is, the higher its degree of strength. It is less clear how we should make sense of an action's degree of strength, however. Actions are not more or less intense, for example. Further, there are various types of action and there is no unique scale – e.g., measuring levels of suffering, or pleasure, etc. – such that it can be relevantly applied to all actions. For example, it is not relevant to measure the level of suffering when trying to make sense of how much I recommended Leo Tolstoy's *Anna Karenina* to someone.

It should be mentioned that those favouring the attitudinal account make sense of values' degree of strength strictly by appealing to attitudes' intensity, not actions' degree of strength (whatever that is). I will follow this approach in this thesis. In Chapter 3 (§3.6.3), I will say more about what makes it fitting to blame someone more or less intensely, and in Chapter 6, I will get back to the issue that actions do not come in degrees in a way that is helpful for FA scholars that want to make sense of a certain object's degree of value.

¹⁶ For a defence of making sense of values' degree in the way I prefer, see Andersson and Werkmäster (2020).

Now to the challenge, human beings are limited in the sense that they cannot feel or believe things to just any degree of strength – there is a limit to how intense our attitudes can be. However, there is no corresponding limit with regards to values. Some objects might even be claimed to be infinitely good. It is thus unclear whether we can rely on the attitudinal account to make sense of infinitely good objects, as it seems impossible for us to favour something to an infinite degree.

I do not pretend to have an adequate answer to the above worry. Tentatively, I believe that defenders of the attitudinal account do have the resources to account for infinitely good or bad values. For example, they might reply that an object is infinitely good (granting there are such objects) if, and only if, it is fitting to favour it maximally (or, perhaps, more than what one is currently doing). That said, it may nonetheless be the case that we have fewer degrees of (dis)favouring than there are degrees of value, but that is a question for another day.¹⁷

I suggest, therefore, we can improve *The FA analysis (generic)* as follows:

The FA analysis: An object x is (dis)valuable to degree n if, and only if, (and because) it is fitting for anyone to (dis)favour x to degree n .

The FA analysis of value, thus stated, is the one I will have in mind in what follows.

1.2.1 FA Blameworthiness

As stated, in this thesis I aim to formulate an informative FA analysis of blameworthiness and test it by appealing to some central claims made in the debate about the ethics of blaming. As a first pass, we can formulate the FA analysis of blameworthiness as follows:

FA blameworthiness: an agent A is blameworthy to degree n for having ϕ -ed if, and only if, (and because) it is fitting for anyone to blame A for having ϕ -ed to degree n .

¹⁷ Crisp (2005) and Reisner (2009) suggest respectively that we also have fewer attitudes than value properties. Consider, for instance, what Crisp writes here: “Consider, say, grace and delicacy. To be sure, they call for certain responses, but the responses themselves are too similar to enable us to distinguish the properties” (Crisp 2005: 82).

To make this analysis more informative, we need to say more about the attitudinal part of it – i.e., about what is meant by “blame” – and what makes it fitting for anyone to blame A for having ϕ -ed to degree n . These are first-order questions that I aim to answer in Chapters 2, 3 and 6.

In Chapter 2, I will show that there are various accounts of the nature of blame. Proposals include the suggestion that to blame is to make a certain judgment and the idea that to blame is to be angry with someone, to name just two. Not only that, but in the pertaining Chapter 2, I will also argue that blame can be kept private, i.e., need not be communicated to anyone at all, and can also be (sincerely) overt, as happens when the agent acts overtly on, for example, the desires her blame prompts.

In Chapter 6, I will argue that we should take “blame”, in the FA analysis of blameworthiness, to refer to just private blame. Otherwise, roughly, the analysis will be in tension with central claims often made in the debate over the ethics of blaming.

Finally, in Chapter 3, I will defend the view that what makes it fitting for anyone to blame an agent for having ϕ -ed to degree n is the fact that she deliberately performed an action, or omitted to act, despite being aware of the facts that constitute normative reasons which, taken together, make the act or omission wrong.

The locution “performed an action, or omitted to act” is cumbersome and adds to the reader’s cognitive load. In this thesis, I will therefore often use the shorthand “performed an action or omission” instead. I recognise that linguistically this is not ideal, since omissions are not performed. However, I think my meaning is clear, and that the shorthand renders the sentence in which it appears less prolix and easier to follow.

1.2.2 The WKR Problem

The FA analysis of value is claimed to have several favourable features. For example, it is meta-ethically and substantially neutral. Both cognitivists and non-cognitivists, utilitarians and Kantians can accept the FA analysis of value.

Further, the FA analysis of value has proved theoretically useful when making sense of various values, including final value, instrumental value and personal

value. In short, and taking inspiration from Toni Rønnow-Rasmussen (2011), we can distinguish between these different kinds of value by changing the attitudinal component:

Final value: an object x is finally valuable to degree n if, and only if, (and because) it is fitting for anyone to favour x *for its own sake* to degree n .

Personal value: an object x is valuable for me to degree n if, and only if, (and because) it is fitting for anyone to favour x *for my sake* to degree n .

Instrumental value: an object x is instrumentally valuable to degree n if, and only if, it is fitting for anyone to favour x *for the sake of its effects* to degree n .

Finally, the FA analysis of value helps us to demystify the notion of goodness with the claim that being good just is being something what there is sufficient reason for anyone to favour.¹⁸

That said, there are several pressing objections to the FA analysis of value. It is claimed, for example, that the FA analysis of value is unable to account for valuable objects in worlds where there are no agents. In such worlds, it cannot be fitting for anyone to (dis)favour certain objects – because there are no agents – but this does not show that there is nothing of value in those worlds. This is called the *solitary goods objection*.¹⁹

In addition, the FA analysis of value is criticised for not accounting properly for cases where, for example, two objects seem equally good – hence, it should be fitting for anyone to favour them equally – but where for one agent it does not seem fitting for her to value the objects equally owing (say) to personal ties with one of the objects. This is called *the partiality problem*.²⁰

The solitary good objection and the partiality problem are important challenges. An adequate FA analysis of value has to answer them in a satisfactory way. That said, it is not my aim in this thesis to provide such a satisfactory answer: this is not a defence of the FA analysis of value. Rather, I want to formulate an

¹⁸ For criticism of the claim that an appeal to reasons would be less mysterious, see Olson (2018).

¹⁹ For more on this objection, see Bykvist (2009), Orsi (2013), and Reisner (2015).

²⁰ For more on the partiality problem, see Lemos (2011), Olson (2009), and Zimmerman (2011b).

informative FA analysis of blameworthiness and test it by appealing to some central claims made in the debate about the ethics of blaming. Therefore, in what follows, I will only focus on an objection that I will discuss in this thesis (especially in Chapter 6), namely the *wrong kind of reason problem* (WKR problem). This states that the FA analysis of value is extensionally inadequate. In some cases, we have a sufficient reason to favour something that clearly has no value. In others, we have a sufficient reason not to favour something that clearly has value. An instructive illustration of the WKR problem is Roger Crisp's (2000) saucer-of-mud example: an evil demon threatens you with severe suffering unless you favour a saucer of mud. Consequently, you have sufficient reason to favour the saucer of mud. However, this does not show that the saucer of mud is valuable. Therefore, the FA analysis of value must be false.²¹ In Chapter 6 (§6.2), I argue that the challenges from the ethics of blaming to the FA analysis of blameworthiness are variants of the WKR problem.

The most common way of responding to the WKR problem is by distinguishing between *the right kind of reason* (RKR) and *the wrong kind of reason* (WKR) to (dis)favour, and then to claim that the FA analysis of value is only concerned with RKR. In short, the reason appealing to the demon's threat is of the wrong kind, therefore, the saucer of mud is not valuable – only RKRs determine value. We can reformulate the FA analysis of value to accommodate this manoeuvre like this:

The FA analysis (RKR): An object x is (dis)valuable to degree n if, and only if, (and because) there is a sufficient RKR for anyone to (dis)favour x to degree n .

Attempted explanations of how, exactly, to spell out in detail the distinction between the notion of an RKR and that of a WKR have now become a small cottage industry of their own. No consensus has so far emerged. Briefly, the proposals in the literature either face counter-examples or turn out to be circular.²²

²¹ In a trivial sense, the demon's threat gives you sufficient reason to favour the saucer of mud as a means of avoiding suffering, thereby making the saucer of mud instrumentally valuable. A harder case arises if the demon threatens you with suffering unless you favour the saucer of mud for its own sake.

²² For presentation and assessment of some of the various proposals in the literature, see D'Arms and Jacobson (2000b), Danielsson and Olson (2007), Gertken and Kiesewetter (2017), Hieronymi (2005), Parfit (2011), Rabinowicz and Rønnow-Rasmussen (2006), Schroeder (2010; 2012), and Sharadin (2016).

That said, in Chapter 6, I discuss a hall-mark that RKR and WKR are thought to have and a suggestion as to how to spell out the distinction between RKR and WKR. I shall now present that hallmark and distinction.

Some argue that there is a *motivational asymmetry* between RKR and WKR.²³ The crucial thought here is that it is difficult, or even impossible, to be motivated by WKR but not RKR. It is difficult, or even impossible, to admire a saucer of mud just because a demon has threatened one, demanding that one admires it.

According to another suggestion, RKR are reasons that are shared by everyone engaged in a certain activity, because they are engaged in that activity. WKR, on the other hand, are reasons that are not shared by everyone engaged in a certain activity, because they are engaged in that activity. For example, everyone engaged in the activity of playing chess wants to win (let us suppose). Thus, for everyone engaged in that activity, the fact that a certain move would help them to win is an RKR to make that move partly because they are engaged in that activity. An incentive, offered by a demon, say, to make a losing chess move does not give every other chess-player not so threatened a reason to make that move.²⁴

The other way to respond to the WKR problem is to dissolve it by denying that it is a problem to begin with. According to this strategy, although there *seem* to be a sufficient reason for anyone to favour the saucer of mud given the demonic threat, this is not actually the case. There is, in fact, no sufficient reason to favour the saucer of mud. What there is, at best, is a sufficient reason for anyone to *get themselves to favour it*.²⁵

The main objection to dissolving the WKR problem in this way is, in crude terms, that dissolvers will find it difficult to give a convincing explanation of why the demonic threat only provides you with a reason to get yourself to have an attitude rather than a reason to have the attitude.²⁶ The intuition is that if you have a

²³ See, for example, Heuer (2010).

²⁴ For a defence of this suggestion, see Schroeder (2010; 2012). One can call reasons that are shared by everyone engaged in a certain activity *shared reasons* and those that are not shared by everyone engaged in a certain activity *idiosyncratic reason* (Gertken & Kiesewetter 2016).

²⁵ For defenders of dissolving the WKR problem, see in particular Skorupski (2010a), Cosker-Rowland (2019) and Gibbard (1990).

²⁶ See, for example., Danielsson and Olson (2007), McHugh and Way (2016), and Rabinowicz and Rønnow-Rasmussen (2004).

sufficient reason to get yourself to favour the saucer of mud, you thereby have a sufficient reason to favour it. Dissolvers need to provide principled grounds for thinking that the demon's threat is not really a sufficient reason to admire the saucer of mud, and this is something that it has proved hard to do.

My preferred way of combining the FA analysis of blameworthiness with some central claims made in the debate about the ethics of blaming, presented in Chapter 6, is similar in spirit, but importantly not identical to, the way dissolvers "solve" or bypass the WKR problem. Roughly, in Chapter 6, I try to reject the claim that ideas about standing to blame and proportional blame create problems for the FA analysis of blameworthiness, and I believe I can provide a plausible explanation of why they do not.²⁷

To conclude this long section on the FA analysis of value, the main ideas we need to hold in our minds as background information in the chapters to come are the following: (i) "fitting" denotes the presence of a sufficient reason; (ii) the attitudinal part of the FA analysis of value need not refer to an attitude – it can also refer to an action or some combination of attitudes and actions; (iii) we determine an object's degree of value by appealing to how intense favouring (an attitude) it is fitting to direct upon the object; and finally, (iv) common ways of solving the WKR problem include differentiating between the RKR and the WKR, and dissolving the WKR problem by denying that it is a problem.

²⁷ There is yet another way to respond to the WKR problem. It has gained prominence recently, but for reasons of relevance, I will not examine it in this thesis. This is to abandon the buck-passing version of the FA analysis of value and instead formulate the FA analysis of value so that it is, allegedly, immune to the WKR problem. More precisely, instead of taking "fitting" to mean "sufficient reason", as I do in this thesis, we could take a "fittingness-first approach" and take "fittingness" as basic and understand value in terms of fitting attitudes or responses. Understanding the FA analysis of value in this way, it is argued, avoids the WKR problem. Consider, for example, what McHugh and Way (2022) write here: "[...] fittingness is insensitive to incentives. Thus the demon's threat to destroy the world unless you admire him does not make it fitting to admire him – it does not make him worthy of admiration. So the wrong kind of reason problem doesn't arise for the fittingness-attitude account" (McHugh & Way 2022: 107). This way of solving the WKR problem comes with a cost, however. Rabinowicz and Rønnow-Rasmussen (2004) argue that taking fittingness to be basic involves mystifying the notion of value, which we do not want to do. So, this approach to the WKR problem is not perfect either.

1.3 Methodology and Limitations

In this section, I make some brief general remarks about my methodology and acknowledge some relevant limitations of it.

1.3.1 Methodology

In this thesis, when I analyse concepts or theories, I will often appeal to our *intuitions* about cases. When I write “our intuitions”, “intuitively”, or such, I mostly mean my own intuitions. Of course, I hope and believe that my peers share those intuitions.²⁸

Throughout the thesis, I treat intuitions as providing *pro tanto* evidence for or against certain views about blame. In analytic philosophy, and normative ethics, and in value theory in particular, it is common to treat intuitions in this way. Since I theorise about value theory and normative ethics, I make the relatively uncontroversial background assumption that it makes sense to treat intuitions as providing *pro tanto* evidence for or against certain views about blame.

Further, when I analyse accounts, I sometimes appeal to what sounds “possible”. By this I usually mean conceptually possible unless I state otherwise. When I write, for instance, that a phenomenon sounds possible, I do not mean that the phenomenon is common or that all of us have encountered it. Rather, I mean something weaker. I mean that we do not sound confused when we refer to it.

Finally, the general methodology I use to reach my conclusions is this: I weigh my intuitions, for example, about certain cases, against theoretical considerations about what we want from our accounts or analyses. For example, we want our accounts to be informative. I assume that a good account or analysis is not subject to counterexamples, is intuitive, informative, non-circular, “and so on”. In Chapter 2, I illustrate how these kinds of consideration can come into conflict. There I illustrate that even when an analysis of the nature of blame manages to account for all of our intuitions about blame, that does not necessarily make it a good analysis of the nature of blame. It may still remain uninformative, for example.

²⁸ That is to say, I am not particularly interested in what “everyone” thinks, statistically speaking.

1.3.2 Some General Limitations

In this section, I state some general limitations of the account proposed here. In the chapters to come, I will state some particular limitations as well.

One might wonder why the topic of this thesis is blame and not, for example, *praise and blame*. It is common among both lay people and blame scholars to treat blame and praise as opposites. Therefore, one might ask why I have chosen to focus only on the negative aspect and not both aspects. Two main reasons motivate my narrow focus.

First, blame seems to be a much more serious affair than praise. While many people find it unfair, or harsh, to blame people unfittingly, most of us do not find unfitting praise equally unfairly – or unfair at all. Thus, for practical reasons, it seems more important to investigate blame.

Second, ultimately it is not evident that praise and blame are opposites. In the accompanying Appendix, I will set out some similarities and differences between praise and blame. I will argue that the differences suggest that praise merits its own treatment, and that we cannot assume that what we have said about blame easily translates, in an opposing manner, to praise.

As claimed in §1.2, the ethics of blame incorporates conditions that need to be considered when we are deciding whether, in a particular case, blaming is permissible. I will discuss two such conditions in, respectively, Chapters 4 and 5 of this thesis – *standing* and *proportionality*. But it might be asked why I do not spend time examining other conditions. For example, one important condition seems to be that the blamer is *epistemically familiar* with the blameworthy agent and her act. That is, when an agent is blameworthy for having performed a particular act or omission, it may still not be permissible for you to blame her, because you lack sufficient evidence about her blameworthiness. Another relevant-looking condition concerns what occurs *after* an agent has performed the blameworthy act or omission. Specifically, confronted by an agent who is indeed blameworthy for having performed a particular act or omission, you may *not* be permitted to blame that person if she has blamed herself to a sufficient degree.

These are interesting debates. However, I have chosen to focus on the two conditions mentioned alone, because there are some important problems to solve with respect to those two conditions, and I do not have the space here to investigate all of the conditions that we know to be significant factors in the ethics of blaming.

1.4 Looking Ahead

In this final section, I outline the structure of the thesis and highlight particular research gaps and problems that it aims to fill and solve.

In the next chapter, Chapter 2, I aim to identify an adequate analysis of the nature of blame. I start with some preliminaries. For example, present some intuitions about blaming that any adequate analysis of the nature of blame should account for. Then I present the sceptical claim that it is impossible to settle on a unitary characterisation of the nature of blame because the phenomenon is too diverse. After that, I ask whether current non-sceptical accounts of the nature of blame succeed in accounting for the desiderata. I evaluate views according to which blame is tantamount to: (i) making a certain judgment, (ii) performing some overt act, (iii) being angry with someone, (iv) revising one's expectations towards someone, (v) desiring that the agent had not acted as she did, (vi) having a particular function, (vii) undergoing a certain process. The chapter concludes that none of the existing or possible views listed under (i)-(vii) are satisfactory: all are either susceptible to serious counterexamples or fail to account for the desiderata. Consequently, the search for an alternative account is warranted, and so, I propose a new theory of blame.

The theory I develop states that blame is essentially a sentiment. This means that when we blame someone, we are, in various circumstances, disposed to feel a range of different emotions, think various thoughts, and desire to act in a range of different ways, in response to the blameworthy agent. I argue that the view that blame is a sentiment accounts for the desiderata in an adequate manner, and also manages to be illuminating – and that, therefore, it entitles us to resist the sceptical claim.²⁹

In Chapter 3, I aim to find a plausible account of what makes it fitting for anyone to blame a certain agent for having performed a certain action or omission. As in

²⁹ The chapter develops a view on the nature of blame which I have published elsewhere Johansson Werkmäster (2022). One vital difference between this thesis and the paper is that, in this thesis, I do not claim that blame manifests itself in actions (I claim that in the paper), but rather in desires to act. Apart from this difference, in the paper I *briefly* discuss three accounts of the nature of blame: the judgment account, the act account and the anger account. In this thesis, I discuss these accounts more thoroughly and examine several other accounts of blame (such as Scanlon's account of blame and the functionalist account of blame) as well.

Chapter 2, here I start with some preliminaries. For example, I present the distinction between *de re* and *de dicto* concerns, and I elaborate two cases that any plausible account of what makes agents blameworthy should account for adequately. I then scrutinise two influential accounts of what makes agents blameworthy for having performed particular actions or omissions – the *witting wrongdoing view* and the *wrong-making view*. I ask how well they account for the two cases. My argument is that both have a hard time accounting for the cases adequately. Eventually, I provide a novel modification of the wrong-making view and defend it briefly. I argue that the modified wrong-making view, apart from being able to make sense of the cases, can also explain other core claims that are made, and issues that arise, in the debate over what makes agents blameworthy. Thus, it explains the claim that we can be blameworthy to various degrees, and the claim that blameworthiness and wrongness are distinct concepts that should not be merged together. The results arrived at in Chapters 2 and 3 are later used, in Chapter 6, to illuminate and improve the FA analysis of blameworthiness presented in §1.2.

In Chapter 4, I argue, contrary to the common opinion, for scepticism about the claim that we need to have standing in order to blame someone privately. Briefly, I contend that key features of standing, such as the features that conditions of standing guide blamers and the feature that standing is a normative power, do not apply to private blame in a way that justifies us in thinking that we need a standing to blame someone privately.

In Chapter 5, I explore what it is for blame to be proportionate. Nearly nothing has been written about proportional blame. I argue that given overt and private blame's differing natures, and especially the way in which they aggregate differently, private and overt blame should be associated with different proportionality principles. I also present some practical implications of my account.³⁰

³⁰ Chapters 4 and 5 are built upon the papers "Scepticism About the Idea of Standing to Blame Someone Privately" (under review) and "Blame and Proportionality" (revise-and-resubmit), which were co-written with Jakob Werkmäster. It is important to note that the papers are *co-authored* – i.e., we have made equal contributions to the papers, especially to the main arguments. Another thing worth mentioning is that while my claims about standing to blame and proportional blame in this thesis are firmly grounded in my accounts of the nature of blame and of what makes agents blameworthy, our claims in the papers are not. Finally, I consider arguments and claims in Chapters 4 and 5 that we do not consider in the papers (e.g.,

In Chapter 6, I test the FA analysis of blameworthiness by relating it to central claims made in the debate about the ethics of blaming. Drawing on insights from Chapter 4 and 5, I argue that the ideas of proportional blame and standing to blame might be in tension with the FA analysis of blameworthiness if it analyses blameworthiness by appeal to fitting overt blame. My preferred way of resolving the tension is by restricting the FA analysis of value, including the FA analysis of blameworthiness, so that it analyses values solely in terms of fitting (private) attitudes.

In Chapter 7, I provide a summary of the main results and identify directions for future research.

The thesis also includes an Appendix where I investigate whether we can generalise my conclusions about blame to praise. As stated in §1.3.2, it is commonly believed that blame and praise are, more or less, opposite sides of the same coin. Therefore, it might be considered to be an advantage if my account of blame easily translates to praise. I argue, however, that while we can generalise some conclusions fairly easily, we cannot easily generalise them all – there are some important asymmetries between praise and blame.

the claim that hypocritical private blame is wrong because we sometimes feel guilt when we blame someone privately and hypocritically).

2 On the Nature of Blame

In Chapter 1 (§1.2) I stated that for the FA analysis of blameworthiness to be illuminating, we need to know more about what the term “blame” stands for in that analysis, and about what makes someone a fitting target of blame for having performed a particular act or omission. The task of this chapter is to illuminate the nature of blame. The following chapter aims to carve out what makes someone a fitting target of blame for having performed a certain act or omission. In Chapter 6, I will use the results I arrive at in this chapter when arguing for my preferred interpretation of “blame” in the FA analysis of blameworthiness.

My strategy is as follows. I start with some preliminaries. In §2.1, I clarify the question I aim to answer in this chapter and present some important intuitions and desiderata for which an adequate account of blame needs to account. After this, in §2.2, I relate my discussion in this chapter to common distinctions that are made between moral blame, non-moral blame, self-blame, other-blame, epistemic blame, and similar phenomena. I also, in §2.3, present and discuss the sceptical claim that it is impossible to settle on a unitary characterisation of blame since the phenomenon is too diverse. Having clarified these matters, I move on, in §2.4, to argue that blame is not to be identified solely with one feature, such as a judgment, overt action, or an angry emotion. Then, in §2.5, I discuss and reject two accounts of the nature of blame, offered by George Sher and Thomas M. Scanlon, respectively. These identify blame with more than one feature – e.g., with a judgment and the modification of one’s relation to the wrongdoer. After evaluating and rejecting those accounts, in §2.6-§2.7, I consider some other accounts of blame that, at least at first sight, seem to have the tools to account for the intuitions presented in §2.1, and which also avoid the objections that led us to reject the views discussed in §2.4-§2.5. These accounts state that blame either ought to be identified with a certain function or is to undergo a certain process. I argue against both. Thus, it turns out that almost all existing and possible accounts of the nature of blame have serious flaws. Given this, I develop my own view, that

blame is a sentiment, in §2.8. I show that this view avoids objections to the previous accounts. In that section, I also discuss whether the view that blame is a sentiment implies that blame is a harmful, harsh, or very negative attitude. If it is, we would need very strong reasons for blaming, just as we need very good reasons to punishing someone, given punishment's harsh nature. I end, in §2.9, by presenting an answer to the sceptical claim presented in §2.3 – i.e., that there cannot be an illuminating and unifying analysis of blame.

2.1 Clarifying the Aim and Desiderata

I will not be searching for a lexical definition of the nature of blame. That is, I do not wish to clarify how blame is understood in ordinary language, or by lay people. Rather, I am looking for an *analysis* of the nature of blame – one that contains its necessary and jointly sufficient conditions.

A good analysis of the nature of blame should display certain structural features and should not deviate too far from our intuitions about blame. The structural features ought to include the following: the analysis will be clear, informative, and simple, and it will be useful in normative investigations, and improve our understanding worldly phenomena. To ensure that they are explicit, it will help to set out some of the vital intuitions which, in my view, we have about the nature of blame:

So, blame is:

- a) about something;
- b) connected with various different emotions: e.g., anger, disappointment, sadness, guilt;
- c) connected with motivations to act and actions: e.g., motivation to demand an apology or explanation, social distancing, ostracism;
- d) connected with thoughts and evaluations: e.g., the evaluation that a particular agent ought to apologise, the thought that it is good that others blame a particular agent; and
- e) relatively long-lasting, i.e., lasting for several minutes, hours, days or even years, usually until we feel proper amends have been made.

I believe that most of the intuitions above are fairly clear. However, I recognise that there may be concerns about (e). There do seem to be several cases in which blame is short-lived.

To clarify, I do not deny that blame can be short-lived. What I mean by “relatively long-lasting” is that we typically blame people until they have made proper amends or have apologised, and that this often takes some time.

2.2 Moral Blame, Non-Moral Blame, Epistemic Blame, Etc.

Some believe that it is tenable to make a distinction between *moral* and *non-moral* blame. For example, I can be a fitting target of blame for having made a mathematical mistake, and that kind of blame is not moral but rather non-moral. Some add further distinctions to the list. For example, self-blame, other-blame, epistemic blame, aesthetic blame, third-party blame, second-personal blame, and similar phenomena. The worry is that it will be unclear, in this chapter, whether I am concerned with moral blame, non-moral blame, epistemic blame, aesthetic blame, third-personal, second-personal blame, self-blame, or something else.³¹

I am not sure that it makes sense to distinguish between moral and non-moral blame. To my mind, it sounds odd to say that I can be *blamed* for having made a mathematical mistake. To me, it sounds more natural to say that I can be *criticised* for having made a mathematical mistake, and this is not blame. As regards the distinctions between epistemic blame and aesthetic blame, it sounds equally unnatural to me to say that we can be to *blame* for having made an epistemic or aesthetic mistake (whatever that is). As with non-moral blame, it sounds more natural to me to say that I can be *criticised* for having made an epistemic mistake, and this is not blame.

That said, I do believe it makes sense to distinguish between self-blame and other-blame, where self-blame is blame towards oneself and other-blame the blaming of others. Whether it also makes sense to distinguish between second-personal blame

³¹ Thank you, Leonhard Menges and Agnès Sophia Constance Baehni, for encouraging me to say more about these different “types” of blame.

and third-party blame, where second-personal blame is the blaming of others and where you are, yourself, the victim, and third-party blame is the blaming of others where you are not the victim, I am unsure. It should be mentioned, however, that nothing important in this chapter hinges on whether the above distinctions are tenable. In section §2.8, I will say more about this issue and argue that the soundness of the above distinctions does not cause problems for my position.

While some of the accounts of blame that I will discuss below rely explicitly on the distinction between self-blame and other-blame, others do not. In addition, defenders of most of the accounts I will discuss state that they are concerned with moral blame, not non-moral blame.

2.3 The Sceptical Claim

At first glance, my endeavour to develop an analysis of the nature of blame may seem uncontroversial. But actually, it is not. There is a growing trend in current blame research according to which we should not aim to analyse the nature of blame. For example, Martha Nussbaum (2016) writes that “blame” is too vacuous a concept to be fit for conceptual analysis. She writes:

In short, while it is very useful to distinguish these different cases, and while we surely learn a lot from the distinctions that these fine philosophers have introduced, human reactions come in many types, and the word “blame” is very imprecise. Maybe it’s not quite as duplicitous as “privacy”, which covers things that have no common thread at all. But it’s pretty empty and uninformative. (Nussbaum 2016: 260)

Miranda Fricker (2016) similarly argues that the practice of blame is too heterogenous for there to be an illuminating and unifying analysis of it. She writes:

Take blame. Let us assume that there *is* an analysis available. The point is we should not expect any such analysis to be very illuminating, owing to the fact that the practice of blame is significantly disunified, and is therefore likely to have distinctive or otherwise central features that may not be present in all instances. (Fricker 2016: 166, emphasis original)

Taken together, these sceptical claims suggest that our intuitions about blame are too heterogeneous for there to be an illuminating and unifying analysis of it.

One might wonder what would constitute a satisfactory response to Nussbaum and Fricker. Put differently, there is a question about how one could falsify their sceptical claim. They do not offer much guidance. Admittedly, Nussbaum does argue that Angela Smith's (2013) suggestion, according to which blame is a moral protest, "is in many ways attractive, but it appears to purchase inclusiveness at the price of vagueness" (Nussbaum 2016: 259). She continues by arguing that once we specify in detail what moral protest is, we probably end up with a too narrow account of the nature of blame.³² This suggests that we can falsify the sceptical claim if we provide an analysis of the nature of blame that both *unifies* all of our intuitions about the nature of blame and is *informative*, i.e., not too vague or thin.³³

Now, I believe that scepticism of the kind Nussbaum and Fricker articulate is a last resort. So, before accepting it, I want to ensure that no analysis of the nature of blame can fit the bill. Ultimately, I will argue that the scepticism can be resisted.

Next, in §2.4-§2.8, I will consider how accounts of blame fare against the desiderata listed above. I will survey most existing accounts of blame and some potential accounts of blame. Importantly, I will not attempt to refute the accounts I consider. My aim is more modest: to raise sufficient doubt about prominent and possible accounts of blame for the development of an alternative to seem worthwhile. The task of developing that alternative will begin by the end of this chapter (§2.8) and continue in the rest of the thesis.

³² In §2.6 I discuss Smith's account.

³³ Fricker (2016) argues in a similar vein: "Successful analysis delivers the highest-common-denominator set of features of *X*; but where *X* is an internally diverse practice there is a significant risk that the highest common denominator will turn out to be very low, delivering an extremely thin account. In particular, it will not be capable of illuminating how the different forms of the practice are explanatorily related to one another" (Fricker 2016: 166).

2.4 Blame Is Not Solely a Judgment, an Action or Being Angry

It is common to analyse blame in terms of one salient feature. Some identify blaming with a judgment, others with an overt action, and yet others with an emotion, usually one of anger. Below, I will critically discuss each of these views in turn.

It should be noted that it is not always clear whether the authors I will be discussing intend to provide an analysis of blame (as I understand “analysis”) or are rather aiming to do something else. At the risk of being uncharitable to some, I will assume that they intend to provide what I understand to be an analysis of the nature of blame.

2.4.1 The Judgment Account

According to what I call *the judgment account*, to blame someone is to make a certain judgment about her. What the crucial judgment consists in, varies from one author to another. According to Michael J. Zimmerman, the particular judgment is the judgment that an agent has “stained” her “record” (about her lifetime moral worth). He writes:

To praise or blame someone, in this sense, is simply to make a judgment about her moral record, a judgment which may form the basis of, but which is not itself, a “reaction” either in attitude or in some robust form of behavior toward that person. (Zimmerman 2002: 556)

Although the contents of the judgment might differ, advocates of this account all analyse blame in terms of a judgment. I will therefore examine all such accounts together under the same heading. This is because, as we shall see shortly, there is an important objection to all accounts that appeal solely to the notion of judgment, irrespective of the judgment’s content.

One virtue of the judgment account is that it is simple and neat. To blame is simply to judge that an agent is, more or less, blameworthy. Further, it captures the intuition that blame is about something, and that it is connected with thoughts or evaluations. Despite these virtues, I believe the judgment account fails.

Below, I provide a case in support of the possibility of judging an agent to be blameworthy without blaming her.³⁴ If this case is sound, I believe we have a sufficient reason to reject the judgment account.

Now to the case. Wallace (1994) writes that one may

believe that an especially charming colleague who has cheated and lied to you has done something morally wrong, insofar as he has violated a moral obligation not to cheat or lie for personal advantage, and yet you may have trouble working up any resentment or indignation about this case. In a situation of this sort it would perhaps be strange to say that you blame the colleague for what he has done. (Wallace 1994, 76)

According to Wallace, in some circumstances we are not prone to react in any way at all to an agent we judge to be blameworthy – e.g., with feeling resentment, or by feeling disappointed, or by desiring to avoid or confront her. In those circumstances, Wallace holds, and I concur, that it would be odd to claim that you blame the agent.³⁵ Put differently, blame seems to involve more than merely making a certain judgment.

If indeed Wallace's case shows that it is possible to judge someone blameworthy and at the same time deny that one blames her, we have a good reason to refute the judgment account. In order to arrive at an adequate analysis of the nature of blame, we need to understand what it is to blame an agent that is distinct from solely judging her blameworthy.

Samuel Scheffler draws a similar conclusion after investigating what it is to value something. Consider what he writes here:

But the proposal that to value *X* is simply to believe that *X* is valuable is unsatisfactory in any case, for it is not only possible but commonplace to believe that something is valuable without valuing it oneself. There are, for example, many activities that I regard as valuable but which I myself do not value, including, say,

³⁴ For simplicity, I write “judging an agent to be blameworthy” instead of “judging the agent has stained her moral record”. I believe we can replace “judging to be blameworthy” with the original formulations in all of the particular accounts presented and still have the same problem.

³⁵ Similar cases and concerns are presented by Smith (2005), see especially p. 470 and p. 476-7.

folk dancing, bird-watching, and studying Bulgarian history. Indeed, I value only a tiny fraction of the activities that I take to be valuable. (Scheffler 2010, 21)³⁶

In short, Scheffler argues that to value something is not just to believe that something is valuable. We believe many things are valuable without valuing them ourselves. Likewise, in my view: we believe or judge many agents to be blameworthy without blaming them ourselves. Thus, blaming or valuing someone, or something, is different from merely making a certain judgment.³⁷

Quite apart from this worry, the judgment account seems ill-equipped to account for some core intuitions we have about blaming. For example, the intuitions that blame is connected with various emotions, and that we often do something when we blame someone.

Finally, at face value, the judgment account does not appear to be useful in normative investigations where we need to make sense of degrees of blame. (Making the required sense is something I seek to do in Chapter 5.) It seems that we cannot rely on the judgment account to make sense of degrees of blame, as we cannot judge something more or less.

2.4.2 The Act Account

According to *the act account*, to blame is to perform some overt action, like booing, performing a speech act, or punishing the blamed individual. No one, to my knowledge, really defends this as an *analysis* of the nature of blame. Some blame

³⁶ Kubala (2017) argues for the psychological possibility of valuing something without judging or believing it to be valuable. He states “[a]s a matter of psychological fact, it is possible to value something [...] without believing it valuable” (Kubala 2017, 60).

³⁷ Like Kubala (see previous footnote), some argue that it is possible to blame an agent without judging her to be blameworthy (Menges 2017; Pickard 2013; Portmore 2022). It is usual to refer to this phenomenon as *irrational blame* or *recalcitrant blame*. The discussion of recalcitrant blame draws much inspiration from D’Arms work on recalcitrant emotions (D’Arms 2003). The idea here is that judgments and emotions can come apart in the sense that we may experience an emotion despite having made a judgment that seems to conflict with it. In some circumstances we may judge, for example, that a monster in a movie does not pose a danger to us and still fear it, or we may judge that an agent is not loveable, but rather cruel, and still love her. In a similar way, blame scholars have argued that we may sometimes judge an agent *not* to be blameworthy and still blame her. Here, I do not take a stance on the feasibility of this phenomenon. For my purposes, it is sufficient to hold we do not believe that simply judging an agent blameworthy is to blame her.

scholars refer to it, though, in order to set up helpful comparisons. Scanlon, for example, locates his account of blame between the act account and the judgment account. In a moment, I will present and critically discuss his account. Furthermore, some instrumentalists about blame seem to assume the act account when they are discussing the justification of blame.³⁸ Roughly, instrumentalism about blame is the thesis that we ought to blame someone if, and only if, the benefits of blaming her outweigh the costs of blaming her.³⁹

Despite these observations, I believe it is fruitful to discuss the act account as a possible analysis of the nature of blame – not only because I believe that some version of it is what most laypeople ordinarily have in mind, but also because it has a virtue which I believe any adequate analysis of the nature of blame should possess. The virtue is that it acknowledges the fact that we often do something when we blame someone. We usually demand explanations from those who have wronged us, and protest about wrongs done by politicians. Despite this attractive feature, however, I argue that the act account fails. It fails because it overlooks the important phenomenon of *private blame* and seems ill-equipped to account for the intuition that, in its nature, blame is connected with various kinds of emotions.

We can blame “privately”, that is, we can do so without performing any overt action and without interacting with someone. If that is true, blame cannot be identified solely with the performance of an overt action. Consider the following case:

Vaccination: Your friend’s friend refuses to be vaccinated against a severe virus. You feel that it is not your place to publicly criticise, or to say “boo” to, her for refusing to be vaccinated, so you do not do that. But you still feel, or are prone to feel, angry with, or disappointed at, her for refusing to take the vaccine. You also think it would be fitting for those who are close to her to criticise her for not taking the vaccine, or to demand that she explains her conduct, and so forth.

In *vaccination*, it is correct to say that the agent blames her friend’s friend, but false to say that she performs any overt action, like booing or punishing her.

³⁸ For more about instrumentalism about blame, see especially Arneson (2003) and Milam (2021).

³⁹ McKenna (2012) has developed a conversational theory about the nature of blame. On this account, roughly, to blame someone is to engage in a certain conversation with her. I do take up his theory in this section because it is not so clear to me whether he would identify himself as a defender of the act account.

Rather, she keeps her blame to herself. Regardless of whether we accept the details of the case, and regardless of whether we believe her omission is justified, in such a case it is accurate to say that the agent blames her friend's friend. This shows that blame is not just the performance of an overt action.

In reply, one might revise the act account to include "private actions". Doing so, one might argue, avoids the above counterexample. Paulina Sliwa provides the following analogy in response to the challenge from private blame:

For many interpersonal practices, there is an intrapersonal correlate. Chess is an interpersonal practice but I can play chess by ("against") myself. I can do this entirely in my head – no other person or physical prop required. Thus, we should not be surprised to find a private correlate to public blame. We blame others privately by thinking things we would say "out loud" were we to blame them publicly: "How could he have done this?" Or we might imagine ourselves accusing the wrongdoer "How could you?!" (Sliwa 2019: 209)

Even if we can avoid the present challenge by modifying the act account to include private actions, I still believe the account ultimately fails. This is because it seems ill-equipped to account for the intuition that blaming is connected with various emotions. We often do something when we blame someone, but this is not all: we also often feel something.

It is important to note that the idea that blame is not to be identified solely with some action does not mean that blame is completely unconnected with actions. Later, I will argue that blame manifests itself in, for example, desires to perform overt actions. I will also argue, in Chapters 4 and 5, that we perform overt actions when we blame someone overtly.

2.4.3 The Anger Account

The final account I wish to consider in this section holds that to blame others is to be angry with them. More precisely, to blame is to resent the individual we are blaming in cases where we are the victim, and to feel indignant with them if we are bystanders to the wrongdoing. Moreover, to blame oneself is to feel guilt.⁴⁰

⁴⁰ See, for example, Wolf (2011) and Wallace (1994).

Most defenders of this account identify these affective responses more or less explicitly with certain reactive attitudes of the Strawsonian kind. In Chapter 1 (§1.1), I presented Strawson's idea about the reactive attitudes and claimed that they include, among other attitudes, resentment, indignation and guilt.

Although the accounts differ somewhat in detail, most of them hold that the particular affective response, in cases where we blame others, is something involving anger, such as resentment or indignation, and, in cases where we blame ourselves, guilt. In addition, in cases where we blame others, it is said that these affective responses motivate hostile or approach behaviour – e.g., the individual is motivated to make the blameworthy agent suffer or to demand that she explain her behaviour. I will therefore gather the accounts together, follow suit, and call them, collectively, *the anger account*.

Defenders of this account have not said much about the nature of the angry affective attitudes. Leonhard Menges (2017) suggests that an agent may either have an *affective episode* of anger or guilt, or an *affective (single-tracked) disposition* to be angry with someone or feel guilty with oneself.⁴¹ Two properties allow us to distinguish between affective episodes and affective dispositions. First, while affective episodes are usually thought to be short-lived, lasting for seconds, minutes or, in rare cases, hours, affective dispositions are usually thought to be long-lived, lasting for several minutes, hours, days, weeks, months, and even years. Second, while affective episodes are commonly thought to have a phenomenology, a “what-it-is-likeness”, affective dispositions are typically not thought to have a phenomenology (at least, not in themselves). An episode of fear involves the experience of various bodily changes including, for example, increased heart rate and feeling warm. In contrast, a disposition to be afraid of certain objects does not involve bodily feelings. Rather, affective dispositions can be viewed as tendencies to experience certain affective episodes with bodily feelings.⁴²

Furthermore, affective dispositions can be either single-tracked or multi-tracked. By the former, I mean that the disposition prompts just one kind of emotion,

⁴¹ It should be noted that Menges (2017) does not commit himself to the anger account. Rather, he uses anger as a paradigm example of a blame emotion and leaves it open whether other emotions, such as disappointment, sadness and so on, are blame emotions.

⁴² For more about this distinction, see Deonna and Teroni (2011) and Tappolet (2016).

such as anger, and in one type of circumstance. By the latter, I mean that the disposition prompts many kinds of emotion, and in various circumstances.

Finally, Menges (*ibid*) notes that angry emotions need not be identified with a judgment.⁴³

The anger account captures the fact that when we blame others, we are usually angry with them, and when we blame ourselves, we usually feel guilty. Furthermore, it can take on board the insight that blame is about something, as emotions are generally thought to be about something. In addition, it can explain why blame is connected with actions: in general, emotions often motivate us to act.⁴⁴ Despite these virtues, I argue that the anger account is too narrow, and that a modified version of it fails to provide an acceptable analysis of blame.

Sher (2005) does not believe that other-blame always involves something angry or hostile. He writes:

We may, for example, feel no hostility toward the loved one whom we blame for failing to tell a sensitive acquaintance a hard truth, the criminal whom we blame for a burglary we read about in the newspaper, or the historical figure whom we blame for the misdeeds he performed long ago. (Sher 2005: 88)

I share Sher's belief that other-blame is not always hostile or angry. When we reflect on our own lived experiences, we soon find that they support his claim. I believe our own lived experiences support another claim as well: that blame – both other-blame and self-blame – is connected with a range of different emotions (e.g., disappointment, disgust, hurt feelings, etc.) and not just, say, angry or hostile emotions posited in the case of other-blame. A view of the nature of blame strictly couched in terms of angry emotions, or guilt, therefore seems too narrow.

Instead of rejecting the idea that various emotions can be instances of blaming, defenders of the anger account can modify their view to include all emotions we believe can be instances of blaming. Such a view could look like this: to blame

⁴³ For defenders of the general view that emotions are identified with judgments, see (Nussbaum, 2001; Solomon, 1976). For a critique of such views, see (Deonna & Teroni, 2011). If we accept that the angry blame emotions need not be identified with a judgment, this can be interpreted as a virtue because they have the tools to make sense of irrational or recalcitrant blame, see footnote 37.

⁴⁴ For more about this claim and the claim above, see Deonna and Teroni (2011).

oneself is to feel guilt, or shame, or disgust, or sadness, etc., and to blame others is to feel resentment, or indignation, or disappointment, or hurt feelings, or disgust, or contempt, or sadness, etc.

Although the modified anger account seems to fare better, in terms of co-extensionality, than the unmodified anger account, it fails on other grounds. The modified account provides us with a disjunctive list of emotions. These emotions differ in, among other things, their phenomenology and the kinds of behaviour they tend to prompt: resentment generally triggers approach or attack tendencies, guilt usually prompts reparation tendencies, and disappointment generally issues in withdrawal tendencies. Not only are they very different, but we can experience the emotions included in the disjunction without blaming someone. For example, I can feel disgusted by my food. Consequently, we want an explanation of when these emotions count as cases of blaming and, more generally, what ties all the different emotions together to count as instances of blaming. Hanna Pickard writes in a similar vein:

If blame is like an emotion, then which emotion is it like? For, there is no ‘basic’ emotion of blame. Indeed, it seems that blame can be connected to a range of different emotions. Most obviously, these include anger, hate, and resentment. But the range can plausibly be extended to include certain other states that have an affective dimension without being uncontroversially identifiable as emotions, such as, for instance, disappointment, indignation and contempt. Moreover, as expected given this range, blame’s manifestations can be equally various. Alongside punishing, blame can also be manifest in berating, attacking, humiliating, writing off, rejecting, shunning, abandoning, and criticizing, to name but a few behaviours. There is thus a challenge facing the suggestion that blame is like an emotion. The challenge is to unite these various emotions and manifestations thereof into a single account of blame. For, given that they can occur without counting as instances of blame, we must explain what makes them count, when they do, as instances of blame. (Pickard 2013, 622-3)

We can call the above worry the *unity worry*.

In the blame literature, there is no clear, or generally accepted, answer to the unity worry. I will elaborate two possible answers and argue that neither is successful – or, at least, not yet.

One way of unifying the emotions is by arguing that they have the same *behavioural upshot*. Drawing on the work of Gideon Rosen (2015), we can argue that affective episodes, or dispositions, that prompt sanctioning behaviours are instances of blaming.⁴⁵ Drawing on John Skorupski's (2010a) writings, on the other hand, we can argue that affective episodes, or dispositions, that prompt withdrawal of recognition, social distancing, ostracism, are instances of blaming.

Generally, angry emotions tend to prompt approach or sanction tendencies, not withdrawal tendencies. Further, disappointment, guilt, shame and hurt feelings are generally thought to prompt, not approach or sanctioning behaviour, but rather withdrawal or social distancing. Accepting the Rosen or Skorupski inspired way of unifying the blame emotions would therefore probably exclude some of the emotions our ordinary thoughts suggest can be instances of blaming. For that reason, this way of uniting blame-related emotions seems to fail.

Another way to unite the emotions is by claiming they have the same *content*. There are no clear suggestions about what that content is. Below, I elaborate on two suggestions.

If the content is "agent *A* is blameworthy", theoretical problems arise for those who wish to analyse blameworthiness in terms of fitting blame. Remember from Chapter 1 (§1.2), in this thesis my goal is to investigate the prospects of an analysis of blameworthiness in terms of fitting blame. That analysis would turn out to be circular if we were to analyse blameworthiness with reference to an emotion with the content "agent *A* is blameworthy".⁴⁶ In addition, acceptance of this suggestion risks over-intellectualising blame, excluding some people whom we think can blame from being able to blame. For example, young children would be excluded, since it is questionable whether they have the concept of blameworthiness.

Maybe we can avoid these worries if we amend the content to "agent *A* is involved in something wrong without an excuse" or "agent *A* willingly and knowingly threatened something important", or something similar. I am uncertain that content of these kinds will avoid the over-intellectualising problem, as it is

⁴⁵ It should be mentioned that Rosen (2015) does not aim to provide an analysis of blaming. Despite this, I believe what he writes about the nature of blame, and especially the way he unifies the blame emotions, in his 2015 paper is helpful in this section.

⁴⁶ It should be mentioned that some FA analysts defend circular FA analyses of value, see Garcia (2018) and Tappolet (2016).

questionable whether children have the concepts “excuse”, “willingly” and “knowingly”.

Drawing inspiration from Rosen again, we could instead seek to unify the emotions by claiming that they all involve the same “retributive” content – namely, the thought that “agent *A* deserves to suffer (to degree *n*) for having ϕ -ed”.

I doubt that this retributive thought is present in all instances of blaming, especially in emotions such as hurt feelings, sadness or disappointment. Further, it appears to be possible to blame someone without thinking that they deserve to suffer (to degree *n*) for what they have done.

Thus, it seems that we have ended up at a dead end. If we attempt to unify the blame emotions in the ways suggested by Rosen’s and Skorupski’s work, we fail to unite all the emotions that we believe can be instances of blaming. The other way is not yet promising either. For if we accept the modified version of anger account without answering the unity worry in a plausible way, we end up lacking a neat and informative analysis of blame.

Moving to a final worry about both the anger account and the modified anger account, it is doubtful whether it even makes sense to view blame as an affective episode. Intuitively, as we might recall from §2.1, blame is typically relatively long-lasting. Blame typically lasts for several minutes, hours, days, or even years – usually until we feel proper amends have been made. And we do not normally say that we experience a “pang” or “episode” of blame, while we do commonly say that we experience a “pang” or “episode” of anger or fear. Relatedly, we do not stop blaming someone after having experienced a bout of anger or disappointment in, for instance, the presence of the blameworthy agent. Rather, we continue to blame her after that experience, usually until we feel proper amends have been made.

These remarks do not amount to a knock-down objection to the anger account. Some defenders of this view claim that we may have a disposition to be angry, and dispositions are usually taken to be relatively long-lasting and to persist in the agent after their manifestation have disappeared. However, this worry invites defenders of the anger account (and the modified anger account) to explain the nature of blame-emotions more clearly.

2.4.4 Summary

To blame is not merely to make a certain judgment. To blame is not solely to perform some overt action either. Nor is it simply to just feeling angry or guilty. Next, I consider accounts of the nature of blame that identify blaming with more than one feature. Perhaps the flaw of the previous accounts is that they focus on one feature of blame, not a plurality of features.

2.5 Hybrid Accounts

In the previous section, I discussed accounts of blame that analyse its nature in terms of just one feature – e.g., in terms of overtly acting in a certain way or making a certain judgment. In this section, I introduce and discuss two hybrid accounts of the nature of blame. These accounts are “hybrid” because they analyse blame in terms of more than one property. I will first present and discuss Sher’s account, and then turn to Scanlon’s account.

2.5.1 Sher’s Account

Sher does not aim to provide an analysis of blaming. Rather, his goal is to provide an *explication* or *theory* of blame’s nature and function. He writes:

I am seeking neither a conceptual analysis nor a dictionary definition but something more akin to a theory. (Sher 2005: 112)

It may therefore seem uncharitable of me to evaluate his account as an analysis. Despite this, I believe we can learn something fruitful by evaluating his account in that way.⁴⁷

According to Sher, to blame is to (i) believe that an agent has acted wrongly or badly, and (ii) desire that the agent had not acted wrongly or badly. To illustrate, to blame your partner for cheating on you is to believe that she has acted wrongly

⁴⁷ It should be noted that Sher characterises blame in a way showing that blame and morality stand or fall together. Since this chapter is devoted to the question of what blame is, I will set aside Sher’s discussion of how blame and morality are connected. For a detailed critique of Sher’s argument, see Hieronymi (2008).

or badly by cheating, and desire that she had not cheated on you. This belief-desire pair is supposed to explain why certain hostile affective responses, like resentment and indignation, are associated with blame.

Sher's account is ambitious and rich in detail. It is built on his refutation of the judgment account, the act account and the anger account. Further, it accounts for features we believe are essential to blame, such as the feature that blame involves some sort of evaluation. Despite this, I think it fails because we can construct plausible counterexamples to his view.

According to Smith, Sher's account is vulnerable to (at least) two counterexamples. Following Smith, we can call the first of these *Mother of Murderer* and the second *Bill and Monica*. Smith formulates the *Mother of Murderer* counterexample as follows:

Consider, for example, the reactions of a mother whose son has been justly convicted of murder. Assume that she judges that her son is blameworthy for the crime (she does not doubt that he is guilty) and that she strongly desires that he had not committed it. She desires this because she knows that his wrongdoing will ruin the rest of his life, and she is deeply distraught by this fact. Her reactions in this case might well take the form of deep sadness, despair, or pity, and these reactions appear to be justified by her belief and desire. (Smith, 2013: 35)

In *Mother of Murderer*, it is true that the first agent believes that the second agent has acted badly and desires that the second agent had not acted as she did, but false that hostile emotions (e.g., resentment) are elicited in the first agent and that she blames the second agent. Smith holds that the mother's reaction is, not one of blaming, but rather a perfectly understandable response to "loved ones who have behaved badly in ways that we strongly desire they have not" (Smith, 2013: 35).

Let us move on to the other case Smith discusses, namely *Bill and Monica*:

Consider, for example, the attitudes many Republicans had on learning of Bill Clinton's ill-fated dalliance with Monica Lewinsky. I think it is fair to say that great many of these individuals *blamed* Clinton for his behavior (or at least for his lack of candor about his behavior). Is it so clear, however, that all of these individuals desired that Clinton had not behaved badly? To the contrary, I suspect most of them were quite happy to see him do wrong, yet they blamed him all the same. In fact – and this is a sad truth about us – I think it is rather common for us

to relish the missteps of others, yet this in no way inhibits our tendency to blame such individuals for their misdeeds. (Smith, 2013: 35, emphasis original)

Bill and Monica is a case where it is true that the agents blame Bill, and believe that he acted badly, and where relevant affective responses are indeed elicited, but where it is false that the agents desire that he had not acted as he did.

Again, Sher does not intend to provide an analysis of the nature of blame. In fact, he is aware that his account might be subject to counterexamples. In reply to those who press those counterexamples, he writes:

If my aim were either to analyze the concept of blame or to define the word “blame”, then the answer would depend on what we would say about a person who had a desire-belief pair of the relevant kind but lacked the corresponding dispositions. It would depend on whether we would describe *X* as blaming *Y* if *X* believed that *Y* had acted badly, *X* wanted *Y* not to have acted badly, but had a psychology that was so nonstandard that he lacked any of the affective or behavioral dispositions that normally accompany such desires and beliefs. For the record, I do not think we would be willing to describe *X* as blaming *Y* under these conditions; but for present purposes, this does not matter. Instead, what matters is that such cases are so nonstandard that our intuitions about them are simply irrelevant. The intuitions are irrelevant because the focus of our inquiry is neither a word nor a concept, but a phenomenon in the world. (Sher, 2005: 112)

I do not believe the cases presented by Smith involve agents with a “nonstandard” psychology. Rather, I think they involve agents that most of us have encountered at some point – even ourselves.⁴⁸

2.5.2 Scanlon’s Account

Like Sher, Scanlon is not presenting an analysis of the nature of blame. Nor does he aim to provide a theory or explication, as Sher does. Rather, he is offering an *interpretation*. What he means by “interpretation” is not very clear to me. However, Scanlon is aware that his interpretation might in some way be revisionary. That is, it is revisionary insofar as accepting it involves changing our

⁴⁸ Regardless of this, what we can learn from Sher’s account is the importance of clearly stating one’s methodological approach. For example, we evaluate conceptual analyses, essential definitions and family resemblances differently.

minds about some things we previously were inclined to believe about the nature of blame. In addition, he thinks that a good interpretation of the nature of blame, such as his own, can provide a thoroughgoing explanation of some puzzles connected with blame. An example is the puzzle concerning resultant moral luck: we seem to think that two careless drunken drivers are equally blameworthy for driving under the influence of alcohol, yet, at the same time, we think that the one who killed an innocent pedestrian while driving deserves more blame than the one who did not kill anyone while driving.⁴⁹ I will not consider in detail how Scanlon's interpretation of blame explains this puzzle in a satisfactory way.⁵⁰ Rather, I am concerned with the question whether his account of the nature of blame is preferable to those we have discussed so far, and whether it is, overall, plausible.

To be clear, I will evaluate Scanlon's interpretation as an *analysis* of the nature of blame. Some might see this – for reasons that are similar to those suggesting that my evaluation of Sher's account was tendentious – as unfair of me. However, I believe that evaluating Scanlon's account in this way moves our discussion of what blaming essentially involves forward.

Recall from §2.4.2, that Scanlon locates his view of the nature of blame in between the judgment account and the act account – i.e., between the view that to blame is to make a certain judgment and the view that to blame is to perform some action, like booing or punishing someone. More precisely, according to Scanlon, to blame someone is (i) to judge that she has impaired her relation with you, and, as a consequence (ii) to modify your understanding of your relationship with her (e.g., to withhold expectations that that relationship would normally involve) in

⁴⁹ Moral luck occurs when factors beyond an agent's control affect her moral status. In the present context, the moral status in question concerns her degree of blameworthiness. Resultant moral luck occurs when an agent performs an action or omission with a consequence that is at least partially beyond her control and that consequence positively affects her moral status. If the driver who kills someone were more blameworthy than the one who does not, this would be a case of resultant moral luck. If they were equally blameworthy, by contrast, this would not be a case of resultant moral luck. In Chapter 3 (§3.6.3), I briefly argue against resultant moral luck. For more on moral luck, see Nagel (1979), Williams (1981), and Hartman (2019).

⁵⁰ In short, he explains the difference by appeal to his ideas about relations and the significance of actions. The individual who killed a pedestrian has impaired relations in a different way than the individual who does not – see Scanlon (2013: 91).

the particular ways that the judgment about blameworthiness in (i) makes appropriate.

We impair our relations to other agents, Scanlon holds, when we act, or hold attitudes, towards them in ways that are ruled out by the standards of those relationships. More precisely, he writes:

Impairment of the kind I refer to occurs when one party, while standing in the relevant relation to another person, holds attitudes toward that person that are ruled out by the standards of that relationship, thus making it appropriate for the other party to have attitudes other than those that the relationship normally invokes. (Scanlon 2008: 35)

The modification that is appropriate depends on how one is related to the blameworthy agent, and to the significance of her action, according to Scanlon. Thus, the appropriate modification will differ depending on whether the impairment was brought out by your partner or a stranger, and whether one is a victim of the action or a bystander, and so on.

In short, what is central to Scanlon's account is a shift in attitudes, or a set of modified expectations, towards the blameworthy agent.

There are many reasons to admire Scanlon's account. A crucial one is that he points to an important feature of blaming that we have not yet discussed, namely the revision of attitudes or expectations. It is true that we tend to revise our intentions and attitudes toward those we blame. We might become less willing to trust or cooperate with them, for example. That said, there are also reasons to be sceptical about the Scanlonian account. In this section, I do not examine all those reasons: I pass over worries about the way Scanlon handles self-blame (intuitively, we do not revise our attitudes towards ourselves) and the blaming of strangers (intuitively, we do not have relationships with all people). I will look only at those that are of particular relevance to this chapter.

Scanlon's interpretation of blame sounds unfamiliar to her, says Wolf. In addition, she thinks that it does not fit with how she ordinarily blames. She offers two counterexamples to his account that are meant to illustrate these claims. The first of these purports to show that relationships may be modified without blame. The latter intends to show that relationships need not be modified even if there is

blame. I believe it is helpful to consider them to bring out what I believe is notably missing from Scanlon's account.

I have taken the liberty of reformulating the cases offered by Wolf so that they fit better into my own discussion. I label the first case *Publisher* and the second *Mother-Daughter*. Starting with *Publisher*:

Publisher: Author has promised Publisher to send a draft of her book on a date they have agreed on. Despite this, Author does not send it. They decide on a new date. But, again, Author fails to send it. Author's conduct is repeated several times. Author and Publisher are good friends. So, despite Author being bad at sending in her drafts, and despite Publisher judging that Author has not lived up to the standards of their relationship, and despite Publisher revising her expectations about Author – not believing she will send her drafts on time, and so on – Publisher would not claim that she blames Author.

Publisher is a case where it is true that Publisher judges that Author has impaired their relation and revised her expectations towards her in a way that is appropriate to the judgment, but where it is false that Publisher blames Author. If we apply Scanlon's account to the case, we arrive at the counterintuitive result that Publisher blames Author.

Let us consider the next case:

Mother-Daughter: Daughter borrows clothes from Mother on a regular basis without Mother's permission. When Mother finds out, she is angry with Daughter. Despite this, Mother would not claim that Daughter has impaired their relation, nor would she claim that she has revised her relation to Daughter – they are a very loving family, she claims.

Mother-Daughter is a case where it is true that Mother blames Daughter, by being angry with her, but where it is false that Mother judges that Daughter has impaired their relation and revises her attitudes toward Daughter in a way that is appropriate to the judgment. Scanlon's account fails to characterise Mother's response as blame. Instead, it says that Mother does not blame Daughter.

What is missing in both cases – and what explains why we think the latter but not the former involves blame – is, according to Wolf, angry emotions. According to her, we think that Mother blames Daughter because she is angry with her, and we

think that Publisher does not blame Author because she is not angry with, or even prone to feel any anger in relation to, Author.

Scanlon has replied to one of Wolf's counterexamples. He does not seem to find it troubling. To illustrate, he replies to *Mother-Daughter* as follows:

Wolf mentions the case of a daughter who borrows her mother's clothes so freely that the mother wants to put a lock on her closet door. Contrary to what Wolf says, however, this reaction seems to me to involve a shift in intention as a consequence of seeing one's relationship impaired. The kind of mother-daughter relationship one has reason to want would involve borrowing and lending clothes with pleasure, trusting the other to exercise the proper restraint and care in doing this. Intentions and expectations of the kind Wolf describes, according to which one does not lend things gladly and take pleasure in sharing, but would prefer to put a lock on one's closet, constitute an impairment in my sense. There is a shift in attitude here that amounts to blame in my view whether it is accompanied by righteous anger or only disappointment. (Scanlon 2013: 99)

In short, according to Scanlon, Mother shifts her intentions and attitudes towards Daughter and, hence blames her.

Even if Scanlon is correct in claiming that Mother revises her attitudes towards Daughter and judges that she has impaired their relations, I still believe Wolf is correct in saying that something is missing from Scanlon's account. I believe the missing feature is some kind of (what we might call) *emotional vulnerability*. It is not true that when we blame someone, we just simply judge that she has impaired our relations and revise, or desire to revise, our expectations towards her. We also experience, or are prone to experience, a range of different affective responses towards her. We are prone to feel sad or disappointed when thinking about how her action has changed our relationship, angry when we encounter her, and happy when we hear that other people blame her, and so on. I believe that an account of the nature of blame that does not account for this aspect of blame ignores the full complexity of the nature of blame.

Admittedly, Scanlon does not deny that emotions can be elements of blame. He acknowledges that the judgment about impaired relations can make certain attitudes fitting, such as anger, sadness or disappointment. He just thinks that emotions are not the only response such a judgment can make appropriate. More

generally, he thinks that an account of the nature of blame that focuses exclusively on the emotions is too thin. It misses something. Consider what he writes here:

I do not deny that these attitudinal responses can be appropriate, and that they are elements of blame. But an account of blame that focused only on these elements would be too thin. Blame also involves other modifications or our attitudes toward a person, including changes in our readiness to interact with him or her in specific ways. (Scanlon 2008: 143)

In addition, Scanlon believes that “there can be instances of blame without moral emotions” (Scanlon 2013: 98).

Although Scanlon recognises that a judgment about blameworthiness can make certain emotions appropriate, I am not fully satisfied with his reply. According to him, it is possible to blame someone without experiencing or being prone to experience any emotion at all towards her on the basis of her wrongdoing, as emotions are not always appropriate to the judgment about impaired relations. It is this possibility that I take Wolf to be challenging with the *Publisher* case. Indeed, it sounds odd to claim that we blame someone when we merely judge that she has impaired our relationship and modify our expectations towards her in a way appropriate to the judgment, without feeling, or even being prone to feel, any emotion at all towards her on the basis of her wrongdoing.

2.5.3 Summary

In §2.4-§2.5 above, I analysed accounts that seek to illuminate the nature of blame with reference to one or more features. I argued that all of them fail. Specifically, I pointed out that all of these accounts are either vulnerable to plausible counterexamples or cannot account for some important intuitions. In the next three sections (§2.6-§2.8), I consider other accounts which seem to have the tools to answer these objections. The first of these accounts for blame in terms of its functional role. The second is the view that blame should be identified with a process. Finally, the third holds that blame should be identified with a sentiment. I will argue for the last of these views.

2.6 The Functionalist Account

Instead of identifying blame with a particular attitude (e.g., a judgment or emotion), action (e.g., a certain speech act or punishment) or combination of attitudes and actions, functionalist accounts of blame identify blame in terms of its *functional role*. Roughly, as long as something has a certain and well-circumscribed function, it is blame according to these theorists.

Before I explicate the functionalist account more thoroughly, it will be helpful to say what I mean by the term “function”, in particular when I talk about the function of blame. So, what I mean by “blame’s function” is the effect, or role, blame has in our minds and/or in our social practice – similar to a heart’s function, which is to pump blood in the bodies of certain organisms.

The function of blame need not rise to the level of intention in the blamer. A blamer can secure blame’s function without intending to do that. To illustrate, suppose blame’s role is to scare people. An instance of blame that scared some has fulfilled blame’s function, and it can do that even though the blamer did not intend to scare anyone. Likewise, a function of punishment is deterrence, but individuals who punish do not always have a conscious desire to deter others from committing crimes. In sum, defenders of functional theorists of blame aim to identify the effect, or role, of blame in our social practices and then claim that whatever has that effect, or role, is blame.

There are various explications of the functionalist account of blame.⁵¹ Here are three. Defenders of some version of *the communicative account* hold that blame is a sort of address that is directed towards the wrongdoer.⁵² The function of the address is to get the wrongdoer to sincerely acknowledge her wrongdoing, usually by feeling guilt and expressing that guilt via apologies and efforts to make amends. In contrast, defenders of some versions of *the moral protest account* hold that blame is a sort of moral protest (not address) that is not always directed towards the wrongdoer and indeed can be directed at the moral community at large. The function of blame is moral acknowledgment on the part of the blameworthy agent

⁵¹ To be fair, not all scholars label their views as “functionalist”. However, their positions can be interpreted as functionalist, and indeed they have been referred to as functionalist accounts (e.g., see Wang 2021).

⁵² For more on this account, see Darwall (2006) and Macnamara (2015).

and/or on the part of others in the moral community.⁵³ In this sense, an instance of blame can fulfil blame's function even when the wrongdoer is thick-skinned or is unable to feel guilt – and we should note that this is something blame cannot do if we adopt some version of the communicative account. Finally, defenders of *the signalling account* argue that the function of blame is to signal the norms that one values or is committed to.⁵⁴

The way in which I have presented the communicative account, the moral protest account and the signalling account may give the impression that blame can only have one of these three functions. This need not be true. Blame can have several functions. That is, the function of blame could be to get the wrongdoer to acknowledge her fault, or to affirm one's moral standing, or something similar. Whether we attribute one or several functions to blame, we are bound to have some worries about the functionalist account, so I will say no more about the issue of what, exactly, is the specific function here.

One virtue with the functionalist account is that it has the resources to account for many of the intuitions we have about blaming, as well as being able to meet some objections to the accounts we have previously considered. Take, for example, the unity worry about the anger account in §2.4.3. Defenders of the functionalist account can argue that what makes these emotions blaming emotions is that they have the same function, or functions – e.g., that of signalling one's commitment to certain norms. In addition, on this view an action, a judgment and a privately held emotion can all be cases of blaming as long as they all have the same function.

Turning to the worries, all specifications of blame functionalism ascribe to blame a function that is more apt for actions, something overt, than it is for something privately held, such as a privately held emotion. Some describe blame in terms that excludes privately held emotions. It is hard to see how a private emotion, such as disappointment, could ever have the effect of getting the wrongdoer to acknowledge her wrongdoing, or signalling a commitment to certain norms, since the emotion need not be revealed to anyone. Disappointment, certainly, can be hidden. In addition, protest and address allude to some action, or something expressed. The idea of unexpressed, or privately held, protest or address sounds

⁵³ For more on this account, see Smith (2013) and Talbert (2012).

⁵⁴ For more on this account, see Shoemaker and Vargas (2019).

incoherent or strange.⁵⁵ In sum, the worry is that private blame lacks the function(s) described above. Therefore, according to the functionalist account of blame, it should not count as blame. However, we believe that private blame is blame. So, the functionalist account of blame must be false.

In reply, one might argue that private blame has the above functions. We are hardwired to wear our emotions, desires and thoughts on our faces or – more generally, to manifest them in our behaviour. Our facial and behavioural expressions are clear enough and communicate to the blameworthy agent what we demand of her, and so on. For example, it will be clear, just from the look on my face and the way I am behaving, that I blame you and want you to acknowledge your wrongdoing. Put differently, one can reply by denying that blame ever can be fully private: blame always has some observable manifestation that is able to fulfil blame's function.

If we assume that privately held emotions, and blame more generally, necessarily involve expressions that clearly indicate what we demand from the wrongdoer and such like, it is easy to see how even private blame can be a protest or an address, and have the effects sketched above. However, I believe that private blame does not work like that – at least, not in all, or even most, cases. It is not true that we always wear our private blame on our faces, or express it in our behaviour, in a way that clearly reveals, for example, what we are demanding of others. So, even though it might be true that in some cases we display our private blame in our faces and behaviour, it is hardly true in all cases, and it is unclear how defenders of functionalist accounts would handle private blame in those cases.

Another reply is to claim that the fact that I am motivated to approach you for having cheated on your partner, or angry or disappointed with you for having cheated on her, even though I do not act on the motives or emotions in any way, signals that I am committed to the norm that we should not cheat. That is enough to fulfil blame's function. So, private blame has the above functions, at least the function of signalling one's commitment to certain norms.⁵⁶

⁵⁵ The communicative account also seems to describe blame in a way that excludes third-party blame, i.e., blame that is not directed towards the wrongdoer.

⁵⁶ An alternative reply is to agree with me that private blame fails to *directly* fulfil blame's function, but argue that instances of private blame fulfil blame's function *indirectly*. Private blame consists of some attitude, or attitudes, that are typically expressed in speech acts that fulfil

That may be so. However, even if we accept this reply, the functionalist account faces another challenge. It is unclear how useful the functionalist account is in certain normative investigations into blame. In Chapter 5, I want to account for degrees of blame – i.e., I want to explain how we can blame someone more or less. Appealing to the function of blame is not helpful to settle the question of degrees of blame. Consider an agent that signals that she is very committed to the norm of not lying when she blames someone for having lied, but claims that she only blames the liar moderately – by just being mildly disappointed with her for having lied. This case suggests it is not relevant to look at the function of blame when determining how much one blamed someone. Apart from this, some of blame’s potential functions do not admit of degrees in a clear way. There is a question about how one can get the wrongdoer to sincerely acknowledge her wrongdoing by feeling guilt and expressing it via apologies and amends to various degrees. And it is unclear how one can morally protest someone’s action to various degrees.⁵⁷

In sum, I do not want to deny that blame has a function, or several functions. I believe it has. But as an analysis of the nature of blame, the functionalist account is not, in my view, convincing. For it to be more convincing, it needs to say more about how we should make sense of blame’s degree.⁵⁸

2.7 The Process Account

According to the process account, to blame someone is to undergo a certain process consisting of thoughts, emotions, various kinds of motivation and actions. This process always follows a certain pattern. For example, you always go from stage 1 to stage 2, and so on. When you diverge from that pattern you no longer

blame’s function. So, private blame is not blame “proper” but still has an important connection to blame. Thank you, David Alm, for mentioning this reply to me.

⁵⁷ Menges (forthcoming) has proposed another challenge to functionalist accounts of blame. It states that functionalist accounts of blame fail to make sense of the idea that forgiving people makes it odd to keep blaming them. Due to time and space, I have not considered this challenge.

⁵⁸ Perhaps this suggest that we need to combine the functionalist account with another account of blame, such as the sentiment account of blame that I will develop below – so that we can make sense of blame’s degree.

blame someone, on this view. In addition, when you blame someone, you experience the process you are undergoing as a process.⁵⁹

Processes unfold over time and do not exist as wholes at every moment of their existence. In other words, they occupy time by *extending* in time, and thus have different parts at different times rather than *continuing* to exist, in their entirety, throughout a given period of time. Consider a movie that lasts for an hour. Watching the movie between minutes thirty and forty is not watching the whole movie. It is watching part of it. It makes sense to regret having to miss the last fifteen minutes of the movie. A movie poster, in contrast, does not occupy time in the same way. Suppose it exists for ten years. It does not make equal sense to regret not seeing it the last year.⁶⁰ This is because the poster does not have temporal parts as the movie does: it exists as a whole at every moment of its existence. Therefore, being in its presence at a given time is being in the presence of the whole movie poster.

In the literature on emotions, it is quite common to view grief as a process. Grief, Peter Goldie writes,

is a process, and is experienced as a process. It is a kind of process, which, borrowing again from Wittgenstein, I will call a pattern; he said, “‘Grief’ describes a pattern which recurs, with different variations, in the weave of our life’. The pattern has certain features. It includes characteristic thoughts, judgments, feelings, imaginings, actions, expressive actions, habitual actions, and much else besides, unfolding over time, but none of which is essential at any particular time. It involves emotional dispositions as well as particular experiences, and there will be characteristic interactions between these. Describable as grief, or as a grieving, it unfolds over time, and is narratable in ways that I will shortly put forward. (Goldie 2011a: 125-6)

Roughly, when we grieve for something, we experience beliefs, emotions and motivational tendencies, and perform certain actions according to a certain trajectory.⁶¹ To illustrate what the process of grief might look like: we start by

⁵⁹ Thank you, Julien A. Deonna and Fabrice Teroni, for proposing this possible view of blame to me.

⁶⁰ This is consistent with saying that one regrets that one was able to enjoy the whole poster for only nine instead of ten years.

⁶¹ For criticism worth considering about the idea that grief is a process, see Marušić (2018).

denying what has happened, then we feel angry about it, after that we start to bargain, then we become depressed, and finally we end up accepting what has happened.

It is not completely outrageous to believe that blame is a process. Think of all the blame games that take place in close relationships. Those blame games seem to unfold in accordance with a distinct choreography. For example, they may start in rage and then evolve into sadness, and so forth.

One virtue of the process account is that it seems to have the tools to account for all our intuitions about the nature of blame. To illustrate, we can argue that we make a certain evaluation, or judgment, concerning blameworthiness at stage 1. Further, we can argue that various emotions are connected with blame, as we may, for instance, experience different emotions during different stages in the process. At stage 2, I experience anger, at stage 3, disappointment, and so on. Furthermore, the process view can account for our intuition that blame is connected with actions, as we can argue that we perform actions as part of the blaming process. For example, at stage 4, I demand an apology. In addition, it can also account for Scanlon's insight into attitudinal revisions. We can argue that at stage 5 we revise our attitudes towards the blameworthy agent. Additionally, the theory can account for the observation that blame is relatively long-lasting, because processes are commonly thought to last for a relatively long time. For example, we do not grieve over something for minutes. Finally, the process account has the resources to answer the unity worry. We can argue that an instance of disappointment is a case of blame, and not something else, because it forms part of a certain coherent pattern or narrative unique to the process of blaming.

Although the process account undoubtedly has potential, I am uncertain about identifying blame with a process.

It may well be true that some blame games, or instances of blame, seem to evolve according to a distinct choreography, but surely not all such blame games and blamings do. Thus, there is a question about what the distinct "stages" of blame are supposed to be. Do we start by feeling shocked? Angry? What do we feel, do or think after that? To put this in other words, it seems hard to pin down the "stages" in an uncontroversial way.

In addition, while it is fairly common to experience grief as a process, this is not so with blame. We do not normally say that we are “in the process of blaming someone”.

What we can learn from the above discussion is that while the process account has potential, it prompts some worries that need to be addressed if it to be considered as an adequate account of the nature of blame.

2.8 The Sentiment Account

In this section, I argue for the following claims. (i) Blame should be identified with a sentiment. More precisely, blame should be seen as a multi-track disposition that manifests itself in a range of different emotions, beliefs and desires to act in various circumstances. (ii) By viewing blame as a sentiment, we avoid the problems raised earlier in connection with other accounts of blame. Finally, (iii) viewing blame as a sentiment significantly affects other inquiries into it.

2.8.1 Sentiments and Blame

In the literature on emotions, the term “sentiment” has been used in many different ways. Some theorists use it to mark an affective episode, others a disposition.⁶² In this thesis, I follow those who take sentiments to be *multi-track dispositions*.⁶³ Thus, sentiments are dispositions which, across a range of different circumstances, manifest themselves in various emotions, beliefs and desires to act. Paradigmatic examples of sentiments include “love”, “hate”, “like”, “dislike”, and “care”. To appreciate the ways in which sentiments manifest themselves in different ways in various circumstances, consider what Jesse Prinz writes about the sentiments “like” and “dislike” and Goldie about the sentiments “love” and “envy”:

If you like someone, then you experience joy in her presence. But you may also experience amusement when she makes a joke, excitement when anticipating your

⁶² Compare, for instance, how the following philosophers use the term: D’Arms and Jacobson (2000a; 2000b), Deonna and Teroni (2009), Helm (2009), Hume (1738), Naar (2018), Vendrell Ferran (2021), and Wallace (1994).

⁶³ See Ben-Ze’ev (2000), Deonna and Teroni (2009), Naar (2018) and Vendrell Ferran (2021).

next encounter, sadness when you are apart, distress when she is harmed, and so forth. If you dislike someone, you may experience anger, disgust, or contempt in her presence. You may even experience *Schadenfreude* when she falls victim to misfortune. (Prinz 2004: 189)

It is an important feature of [sentiments] that they are not simply dispositions to have a single kind of emotions. For example, your enduring love of your children or parents is not just a disposition to have loving feelings towards them when they are in the offing. It can be expressed in a complex structure of possible responses: delight if they succeed in their endeavours; anger if you hear them insulted behind their backs; fear and concern if you think they might be ill; hope if you think that their illness might have a cure; and so on. Even my envy of Mary's success isn't just a disposition to feel envy; it can be expressed in a feeling of delight when I hear that her latest success looks after all as if it's turning into something of a poisoned chalice; I wouldn't be feeling this delight if I weren't envious of her. (Goldie 2011b: 98)

The idea that blame has multiple manifestations and triggering conditions enjoys some degree of plausibility. It makes sense to say that when you blame someone, you are disposed, in a range of different circumstances, to experience various different emotions, beliefs and desires to act. Characteristic emotions, desires and beliefs include, but are not limited to, the following:

In other-blaming:

Characteristic emotions: anger, disappointment, contempt, sadness, hurt feelings, disgust.

Characteristic desires: desires to demand an excuse, demand an explanation, withdraw from the agent, exclude the agent, harm the agent, make the agent feel guilt.

Characteristic beliefs: the belief that it is fitting for others to blame her, that she ought to take responsibility for her action, that we are no longer on equal terms.

In self-blaming:

Characteristic emotions: guilt, shame, disappointment, anger, sadness, disgust.

Characteristic desires: desires to apologise, repair your wrongdoing, approach the victim, hide from you community.

Characteristic beliefs: the belief that it is fitting for others to blame me, that I ought to take responsibility for my action.

In §2.2, we noted that some blame theorists believe it makes sense to distinguish between epistemic blame, aesthetic blame, and so forth. Although I am sceptical about the claim that it makes sense to make all of these distinctions, if we assume that the distinctions are true, we can differentiate between the forms of blame by adjusting the characteristic manifestations, as I did with self-blame and other-blame above. For example, epistemic blame manifests itself in the desire not to ask the person being blamed for advice on epistemic matters, the belief that others should not ask her for advice on epistemic matters, and so forth. So, although it is true that we should make the distinctions mentioned in §2.2, that need not pose a problem for defenders of the sentiment account. As shown above, we can distinguish between different forms of blame by adjusting the characteristic manifestations.

The blame sentiment is normally triggered when the agent is confronted with the particular blameworthy agent and her action in some way (including cases where the agent in question is oneself). That may occur when the agent is contemplating, hearing about, being reminded about, being confronted by, seeing, or perceiving the particular blameworthy agent and her action. To illustrate all of this: you might experience anger when thinking about the fact that your partner has cheated on you, disgust when seeing that she seems unaffected by the fact that she cheated, sadness when thinking about the fact that your relationship with her has changed because she cheated, think that she owes you an excuse when contemplating the fact that she has cheated and not yet apologised for it, or desire to demand an excuse from her when discussing the fact that she has cheated with her. You would not feel these emotions, think these thoughts, or desire to act in these ways if you did not blame her.

Importantly, sentiments are assumed to remain with the agent even after their manifestations have disappeared. This is a difference, it will be recalled from §2.4.3, between affective dispositions (including sentiments) and affective episodes. For example, I still love my partner after having experienced an affective episode of joy in, say, her presence. As hinted in §2.4.3, I still blame someone

after having experienced an episode of resentment when being reminded about her blameworthy action.

Further, all theorists of sentiment seem to agree that sentiments have the property of intentionality: they are about something, directed towards something. For example, I love my partner, hate the government, and such like. Like other sentiments, blame also has the property of intentionality – i.e., it is about something, it latches on to an object, such as a property out there.

However, to say that blame has the property of intentionality *and add nothing more* would be incomplete. One might wonder what the blame sentiment is really about, given that it has multiple manifestations, and given also that each manifestation seems to be about a different thing. More precisely, each manifestation seems to have different *particular objects* as well as *formal objects*. By “particular object” I mean, what the emotion, belief or desire is about, its object. By “formal object” I mean, roughly, how the particular object is represented or evaluated, as threatening, as a loss, and so forth. For example, my resentment is about my partner cheating, representing cheating as an offense, whereas my sadness is about how our relationship changed as a result of her cheating, representing the change as a loss, and so on.

The question about sentiments’ intentionality is puzzling. However, it is not a special puzzle for my view of blame. It is a general puzzle in the philosophy of sentiments. In general, it is unclear how we should explicate what sentiments are about, given that they are manifested in various different ways. Since my view employs the notion of sentiment, it inherits this problem. But I am hopeful that the most promising general solution will be applicable here.

Let me provide a tentative answer, however. One way of making sense of a given sentiment’s intentionality is by summarising what all its manifestations are about. In the case of cheating mentioned above, all of the manifestations seem to be more or less about a particular agent – in this case, my partner. Not only do all of these manifestations seem to be more or less about a particular agent, but they all seem to depict her in a certain way – namely, as someone who has acted or omitted wrongly or badly. In short, we can say that your blame sentiment is about a particular agent (note: this might be yourself) who is depicted as someone who has done something wrong or bad.

To clarify, by “particular agent” I do not necessarily mean a human being who is alive now or has existed. It could be a fictional character – I blame Anna’s husband Alexei for ruining her life. It could even be an object, an animal, or a group, perhaps as long as we ascribe agential features such as a will to those non-persons, which indeed we sometimes do.⁶⁴ To compare with the sentiment “love”, we sometimes love our cat, our car or dead people. And it seems that, likewise, we sometimes blame our dog or a dead person for having behaved badly.

Finally, sentiments, in contrast with affective episodes, are relatively long-lasting – i.e., they last for several minutes, hours, days, weeks, or even years. As we learned in §2.1, blame typically lasts for a relatively long period of time: several minutes, hours, days, weeks and even years. Usually, we continue to blame until we feel proper amends and apologies have been offered.

Some might worry that the sentiment account cannot make sense of short-lived blame, as it takes blame to be a sort of disposition and dispositions generally last for a long period of time.

It is uncontroversial, however, that you can acquire a disposition and then let go of it quickly. Granting that beliefs are dispositions, upon learning that a war has broken out in a country, you start to believe that a war has broken out there. However, shortly after being presented with that evidence, you learn that the conflicting parties managed to resolve their conflict quickly and the country is no longer at war. In that kind of case, you acquire a belief and then free yourself from it quickly in view of evidence to the contrary. I believe similar claims can be made about blame. In some cases, you start to blame someone after witnessing them deliberately committing an action or omission that was wrong. However, shortly after witnessing that, the wrongdoer makes proper amends and you start to let go of your blame towards her, quickly succeeding in that.

Another clarificatory remark concerns unmanifested blame. The worry is that it is unintuitive to claim that we are blaming someone when the disposition of blame has never been realised. In order for it to be accurate to claim that one is blaming someone, the sentiment must have been actualised in some way.⁶⁵

⁶⁴ Thank you, Mattias Gunnemyr, for pointing out this fact to me.

⁶⁵ Thank you, Daniel Telech and Björn Petersson, for mentioning this worry to me.

This worry does not deliver a knock-down objection to the sentiment account. The account can fairly easily be rephrased to avoid it (if indeed the worry does articulate a genuinely unintuitive feature) by adding, for example, a clause saying that the sentiment must have been actualised at, at least, some point in time for it to be true to say that the person is blaming someone in this way. However, I am unsure whether we need to modify the sentiment account so.

To begin with, in most, if not all, cases we acquire the blame sentiment upon witnessing (or at least thinking that that is what we have witnessed) a wrong being performed by someone deliberately. The sentiment is triggered and manifested at, at least, some point in time, and we are aware that we harbour it. So, although it is theoretically possible for us to have the blame sentiment (or the love sentiment, etc.) about someone without knowing it, or without it being ever triggered or manifested, such cases are rare.

That said, I believe it can be considered a virtue of the sentiment account that it is capable of making sense of some rare, although possible, cases involving unconscious or unmanifested blame. To illustrate: the sentiment account can make sense of cases where an agent blames someone, but the blame is masked for a long period of time as a result of, say, depression. Consider the following case. An agent acquires the blame sentiment about a particular person upon witnessing her committing a wrong willingly. However, directly upon witnessing that and acquiring the sentiment, she becomes depressed for ten years. After ten years, her blame is realised again. In such a case, we want to say that the blame she experienced after ten years is the same blame as before she got depressed, and that she blamed the person during the entire time, but that her blame was masked by her depression.⁶⁶

⁶⁶ A more difficult case is this: an agent acquires the blame sentiment about a particular agent upon witnessing her committing a wrong willingly. However, directly upon witnessing that and acquiring the sentiment, she gets depressed for the rest of her life. So, the blame is never realised again. In such a case, we want to say, on the one hand, that she blamed the agent throughout her life, but that the blame was masked by her depression and, on the other, that she stopped blaming the agent because she stopped caring about her being blameworthy due to her life-long depression.

2.8.2 How the Sentiment Account Avoids Previous Objections

I will now explain how viewing blame as a type of sentiment enables us to avoid the objections discussed in previous sections.

Recall from §2.4.1 that the problem for the judgment account was that it failed to capture the fact that we react in certain characteristic ways when we blame someone. We do not merely judge someone to be blameworthy when we blame someone. Similarly, the worry I raised about Scanlon's account in §2.5.2 is that he does not ascribe emotions an essential role in blaming. According to him, we can blame someone without feeling, or being prone to feel, anything at all. Viewing blame as a sentiment avoids these problems, because, on this account, to blame someone is to be prone to feel certain emotions and to desire to act in certain ways with respect to her.

In §2.4.2 I argued that the problem for the act account is that blame is connected with overt actions but cannot be identified with an overt action alone. We can blame someone "privately", without communicating the blame to anyone. On this account, to blame is to be disposed to desire to perform overt actions. In this sense, the account can explain the intuition that blame is connected with overt actions. Further, it can accommodate private blame, since blaming agents need not act on the desires, emotions and beliefs that their blame sentiment prompts. And when the blame sentiment is not triggered, it is also private.

It also avoids the worry about excessive narrowness presented in connection with the anger account in §2.4.3, since blame, on the sentiment account, is a disposition manifested in a plurality of emotions, not just (say) anger. Additionally, it avoids the worry about degrees of blame presented in connection with the functionalist account in §2.6. On the sentiment account, we can make sense of the degree of blame by appealing to the intensity of the emotions it prompts: the more intense emotions it prompts, the more we are blaming someone. In Chapter 5 (§5.1), I will say more about this way of making sense of blame's degree.

What about the unity worry about the anger account that I presented in §2.4.3? How do we determine whether an instance of disappointment is a case of blaming? And how do we explain what ties instances of all the characteristic emotions, beliefs and desires to act together as instances of blaming? Now, apart from the fact that the manifestations seem to be about a particular agent – one depicted as

having done something wrong or bad – and adopted for the same reason (in the next chapter I carve out what makes it fitting for anyone to blame a certain agent), we can say something further. In order to understand that reply better, it is helpful to first say what I take the nature of sentiments, or more generally that of dispositions, to be.

I follow Hichem Naar (2013; 2018) and take sentiments to be actual properties of persons. These properties are connected with, or directed towards, their manifestations but are not to be identified with them. The two are distinct. Consequently, a number of disassociations of them are possible. An agent might in suitable circumstances emote, believe and desire in a way that is characteristic of a person who blames someone without having the disposition that we call blame. Conversely, she might have the disposition we call blame without in suitable circumstances emoting, desiring and believing in a way that is characteristic of a person who blames someone. Further, these properties are taken to play a causal role in producing the manifestations with which they are connected.⁶⁷

Granting such a roughly specified realist conception of sentiments as dispositions, we can provide a further answer to the unity worry. In short, what unites all the manifestations is that they have the same origin – i.e., a certain property which we call blame. An instance of disappointment is an instance of blaming when it is produced by the property that we call blame. In the context of love, Naar (2013) says something similar:

What ties all the events of a given sequence together as constituting an expression of love? The answer, on the dispositional account, is simply that a disposition, love, is their common origin. And we now have a way to tell whether or not a given event is part of the relevant sequence: we need to consider the disposition that produced it; if it is the disposition with which we identify love, then the event is an expression of love, and if it is not, then the event is not an expression of love. What distinguishes an episode of joy towards another person's embarrassment as an expression of love from a similar episode that is an expression of cruelty is thus that the former originates in love while the latter originates in cruelty. (Perhaps certain cases can be interpreted as mimicking cases.) This point is trivial only superficially, however: whatever sort of thing love-the-disposition is – so far, the

⁶⁷ For other realist theories of dispositions, or power accounts, see Molnar (2003) and Heil (2003).

account is silent on this more specific question – an event that counts as an ‘expression of love’ must come from that thing. (Naar 2013, 353)

Likewise, what ties all instances of the characteristic emotions, beliefs and desires to act together as instances of blaming is that they are all produced by the property we call blame.

Relatedly, it seems that some of blame’s characteristic manifestations are also characteristic, as manifestations, of other sentiments. For example, disgust can manifestation of hate, not just blame. How do we decide whether an instance of disgust is an instance of blaming or one of hating, then? The answer is similar to the one I gave above: disgust is an instance of blaming when it is produced by the property that we call blame, and it is an instance of hating when it is produced by the property that we call hate.

These answers to the unity worry will probably not satisfy everyone. Probably, some will complain that one of the ways of uniting the reactions we connect with blame I have presented is uninformative, if not circular – is it not unhelpful to say that an instance of blame is blame when it is produced by a certain property that we call blame?⁶⁸

We ordinarily explain our own and others behaviour and mental states by postulating entities, such as beliefs and sentiments, that we think are causally responsible for their occurrences. For example, Jakob hugging his son Elliot is an expression of love because it was produced by his love for him. Compare this to a case where Jakob hugs his son Elliot because Jakob has taken a drug. In such a case, we want to say that the conduct is not an expression of love because it has the wrong origin, because it was not produced by Jakob’s love for Elliot. Such kinds of explanation are not just ordinary – they also seem informative.

Still, one might want a better unifying explanation, such as those I discussed in §2.4.3. For example, what unites the emotions is that they all have the same retributive thought that “the wrongdoer deserves to suffer (to degree *n*) for having ϕ -ed”. I am uncertain whether we can find better unifying explanations than those I have provided in this section, and I would prefer my analysis of blame to be

⁶⁸ Thank you, Leonhard Menges and Matthew Talbert for pressing me on this point.

accurate rather than simplistic. I leave the task of finding a better unifying explanation for future research.⁶⁹

To conclude, I believe the problem with previous accounts is that they mistakenly identified the nature of blame with certain of its manifestations. This failure is like identifying love with certain manifestations of love. To love is not solely to desire to give roses. It is not only to feel happy in the presence of the beloved. Rather, love is a sentiment. That is, it is a multi-track disposition which, across a range of different circumstances, manifests itself in various emotions, beliefs and desires to act.

We can summarise the sentiment account of the nature of blame as follows:

The sentiment account of the nature of blame:

Self-blame: an agent A blames another agent B for having φ -ed, and where $A=B$, if, and only if, A has acquired a specific disposition β_1 .

Other-blame: an agent A blames another agent B for having φ -ed, and where it is not the case that $A=B$, if, and only if, A has acquired a specific disposition β_2 .

“ β_1 ” and “ β_2 ” are the descriptions of self-blame and other-blame I provided in §2.8.1. For example, in the cases of other-blame, β_2 involves being disposed to resent, feel indignation, desire to demand an apology, and believing that it is fitting for others to blame the agent.

2.8.3 The Sentiment Account and Badness

Recent work on the nature of blame suggests that blame, or some of the expressions of blame, is bad or has harmful effects.⁷⁰ To appreciate some of the effects, consider what Derk Pereboom writes:

⁶⁹ Perhaps we get a better explanation by combining the sentiment account with the functionalist account (§2.6) – e.g., claiming that what unifies the manifestation is that they all signal one’s commitment to certain norms or values. Future research can explore the plausibility of that suggestion.

⁷⁰ When I write “harmful” I mainly mean harmful in a *morally significant* way.

But often expressions of moral anger have harmful effects. They often fail to contribute to the well-being of those to whom they are directed. Frequently expressions of moral anger are intended to cause physical or emotional pain. Partly as a result of these problems, moral anger often has a tendency to damage or destroy relationships. In extreme cases, it can provide motivation to take very harmful and even lethal action against another. (Pereboom 2009: 172)

One might think that although blame does not result in any harmful effects, the fact that it prompts retributive desires, emotions or beliefs indicates that it is an overall bad attitude.

Before concluding this chapter, I want to ask whether the view that blame is a sentiment implies that blame is overall bad or harmful. If blame is, overall, a bad or harmful attitude, then that speaks against blaming. More precisely, we would need very strong reasons to blame. Likewise, given the harshness of legal punishment – i.e., that such punishment involves an intention to cause harm to someone and often does harm the target significantly – we need very strong reasons to legally punish. Thus, how we answer the question above has important implications for the normativity of blaming.

My view that blame is a sentiment does not commit me to the claim that blame only prompts emotions, beliefs or desires that can be viewed as bad. On my view, blame prompts some such emotions, desires and beliefs, but it prompts other emotions, desires or beliefs as well. It also prompts more “civilised” or “neutral” emotions, beliefs and desires, such as disappointment, hurt feelings, the desire to ask for an explanation, and so on. Further, sometimes, acting on some of blame’s manifestations can lead to harmful or bad effects. For example, acting on the desire to withdraw from the blameworthy agent can do damage to a valuable collaboration with her.⁷¹

Consequently, it is not evident that, on my account, blame is an overall bad attitude, an attitude that we would need very strong reasons to be justified in

⁷¹ In Chapter 5 (§5.1), I say more about blame and harm. Briefly, I write there that (sincere) overt blame, i.e., acting on the desires/emotions/beliefs the blame sentiment prompts, often involves harm as a side-effect. But, importantly, in single-blamer scenarios, the harm of overt blame is often not morally significant, or at least not particularly morally significant. In addition, there is no necessary connection between a single instance of overt blame and harm – e.g., the target may find your overt blame amusing. However, in plurality of blamers scenarios, the harm can be morally significant.

adopting. Compare: the mere fact that love sometimes prompts bad emotions, beliefs or desires – jealousy, intense sadness when the loved one is away and a desire to keep the object of love to oneself, for example – and that acting on those manifestations may sometimes lead to bad consequences, does not entail that love is, overall, a bad attitude.

2.9 Conclusion

The guiding question of this chapter was: What is the nature of blame? I identified some intuitions and theoretical desiderata for answering this question, and used these mainly to criticise prominent and possible accounts of blame. I introduced an alternative account, according to which blame is a sentiment. Having clarified how the account is to be understood, I showed that it does well by the desiderata and answers other concerns the prominent and possible accounts of blame face.

I have not yet said anything about how we can rely on the sentiment account to answer the sceptical claim I presented in §2.3. Recall that, the claim is that our intuitions about the nature of blame are too heterogenous for there to be an illuminating and unifying analysis of it.

Is the account of blame as a sentiment an illuminating and unifying analysis? I believe that it is. As shown in the previous section, it unites all our intuitions about blame within a single framework. On the account I develop, blame is about something, is connected with various beliefs, actions and emotions because it manifests itself in various beliefs, desires to act and emotions, and is relatively long-lasting.

As regards the question whether the account is illuminating, I believe the account of blame I have developed in this chapter is sufficiently illuminating to resist the sceptical claims such as those made by Fricker and Nussbaum. It tells us what blame is, and it explains how, by viewing blame as a sentiment, we avoid the problems associated with other popular accounts of blame. Given this, I believe it would be false to say that it is too thin or vague.

The chapters to follow show that the sentiment account is also useful in other normative investigations. In Chapters 4, 5 and 6, I show that we can rely on the sentiment account to make a useful distinction between *private blame* and *overt*

blame. By “private blame”, as noted in §2.8.2, I mean cases of merely harbouring the sentiment of blame. Importantly, such blame does *not* involve interacting with its target. Nor does it involve the performance of some voluntary action. The blame sentiment can manifest itself in a desire to demand an apology from a particular person. However, one needs not act on that desire. In contrast, when we (sincerely) blame someone overtly, we have the blame sentiment and act deliberately on its manifestation – e.g., we intentionally act overtly on the desire to demand an apology. In Chapter 6, I argue that we should understand “blame” in the FA analysis of blameworthiness as referring to private blame alone, not to overt blame.

3 On What Makes Agents Blameworthy

In Chapter 1 (§1.2), I stated that for the FA analysis of blameworthiness to be illuminating, apart from needing to state what “blame” refers to in that analysis, we need to know what makes it fitting for anyone to blame an agent *A* for having ϕ -ed to degree *n*. Or more accurately, given what I have said in Chapter 2 (§2.9), we need to know more about what makes it fitting for anyone to *privately blame* an agent *A* for having ϕ -ed to degree *n*.⁷² The aim of this chapter is to spell this out.

Specifically, I will evaluate two influential accounts of what makes agents blameworthy – one that can be interpreted as endorsing a *de dicto* concern (*the witting wrongdoing view*) and one that can be interpreted as endorsing a *de re* concern (*the wrong-making view*).⁷³ In short, the former view states that an agent is a fitting target of blame to the extent that she willingly performed an action or omission despite believing her action or omission to be wrong, and the latter view states that an agent is a fitting target of blame to the extent she was motivated by the reasons that made her action or omission wrong.⁷⁴ Both accounts preserve, in one way or another, the common-sense idea that Strawson (see §1.1) pointed to – namely, that all blameworthy agents exhibit some poor quality of will, or rather quality of mind, when acting or omitting. At the end of the chapter, I will sketch a novel modification of the wrong-making view and defend it against some worries and claims defenders of the wrong-making view have previously said little about.

⁷² From now on, I will mainly use the term “blame” instead of “private blame”.

⁷³ In §3.2, I present the distinction between a *de re* concern and *de dicto* concern.

⁷⁴ Recall from Chapter 1 (§1.2.1): the locution “performed an action, or omitted to act” is cumbersome and adds to the reader’s cognitive load. In this thesis, I will therefore often use the shorthand “performed an action or omission” instead. I recognise that linguistically this is not ideal, since omissions are not performed. However, I think my meaning is clear, and that the shorthand renders the sentence in which it appears less prolix and easier to follow.

As I hope will be clear from this, I will not analyse views according to which agents are blameworthy for having performed certain actions or omissions when their actions or omissions manifest their morally problematic evaluative judgments, cares, or such like.⁷⁵ I set aside such accounts because it seems to me that they are more concerned with appraisals of one's virtues and vices, or with attributability, than they are with blameworthiness for one's actions or omissions, or with accountability (which is what I am concerned with in this chapter). In the next section (§3.1), I say more about the distinction between accountability and attributability.

The plan is as follows: as in the previous chapter, I start with some preliminaries. In §3.1, I explain how the discussion that follows connects with Gary Watson's famous distinction between *accountability* and *attributability*. Having clarified that, I move on, §3.2, to present the distinction between a *de dicto* concern and *de re* concern, as the witting wrongdoing view echoes the *de dicto* concern and the wrong-making view the *de re* concern. Then, in §3.3, I present two cases that any account of blameworthiness, including those I will discuss in this chapter, needs to account for successfully in order to be plausible. With those desiderata and clarificatory points in mind, I move on, in §3.4 and §3.5 respectively, to investigate whether the witting wrongdoing view and the wrong-making view can account for the cases presented in §3.3 adequately. I argue that both have difficulty accounting for the cases. That said, in §3.5, I modify the wrong-making view and show that thus modified it can indeed make sense of the cases in §3.3. In §3.6, I briefly defend the modified wrong-making view further by showing that, apart from explaining the cases in §3.3, it seems to be able to handle some other important challenges and claims that have not been addressed much by defenders of the wrong-making view. These include the claims that blameworthiness and wrongness are distinct concepts that should not be merged together and the claim that we can be blameworthy to various degrees. I end with a few concluding remarks in §3.7.

In what follows, I assume that the agents I will be talking about are (*morally*) *responsible agents*. That is, they have the capacity to assess (both "moral" and "non-moral") normative reasons for and against certain actions or omissions, to modulate their behaviour in light of their awareness of these reasons, and so on.

⁷⁵ For example, Smith's (2005) view and Talbert's (2017) view.

In addition, I will also assume that they are not in some form of extreme circumstances at the moment of acting or omitting that would hinder their abilities or capacities from working properly, such as suffering from sudden paralysis or being a victim of a major earthquake.

3.1 Attributability, Accountability, Etc.

Some theorists take Watson's distinction between *attributability* and *accountability* as a starting point when formulating their theories of what makes agents blameworthy.⁷⁶ They might complain about me not similarly using that distinction as a starting point in this chapter, or they may feel confused, asking themselves: Is she concerned with *attributability* blameworthiness or *accountability* blameworthiness? Therefore, before moving on, I want to say a few words about how this chapter relates to that distinction and other similar claims.

Two reasons explain why I have not deployed the distinction between *accountability* and *attributability* in this chapter: (i) it is not very clear to me what the distinction is about, (ii) nor is it clear whether my views about the nature of blame map onto it.

There are several different explications of the distinction between *accountability* and *attributability* on offer. That said, most writers seem to agree with *one or more* of the following claims: (i) *accountability* is concerned with whether punishment, sanction or (harsh) overt blame, is fitting, and *attributability* concerns whether milder responses, or forms of blame, are fitting, such as merely making a certain judgment, or feeling sad or disappointed are fitting; (ii) exhibiting *voluntary* control is a necessary condition of being held accountable for one's action or omission but not for an action or omission being attributable to one; and (iii) *accountability* is concerned with blameworthiness for one's *act* or *omission*, while *attributability* is concerned with more than this, such as with blameworthiness for one's attitude, lack of attention, action, omission – basically “anything” insofar it

⁷⁶ For more about this distinction, see Watson (2004). Shoemaker (2015) adds *answerability* to the list. I do not relate my discussion in this chapter to his tripartite distinction because it is not widely accepted.

reveals something about one's character, cares, or such like. So, am I, in this chapter, concerned with accountability or attributability?

It is difficult to answer this question. Here is why. In this chapter, my concern is with when *private blame* is fitting, not when overt blame is fitting, and my understanding of private and overt blame does not fit perfectly with the distinction between harsh treatment and mild treatment I sketched above. On my view, as will be recalled from Chapter 2 (§2.8.3), private blame is not just “mild”. It can also manifest itself in resentment and guilt, and these are commonly thought to be, not so mild, but rather quite harsh, emotions. In addition, on my view to (sincerely) overtly blame someone is to act on the desires, and such, that the blame prompts. The blame sentiment need not manifest itself merely in, say, retributive desires. It can also be manifested in desires to ask the wrongdoer for an explanation, or offer an apology to the victim, and acting on those desires need not amount to harsh treatment.

As regards (ii) and (iii), in this chapter I am concerned with blameworthiness for one's *act* or *omission* and the accounts I will discuss presuppose that one had voluntary control over one's action or omission.

Given these points, it might be claimed that, in this chapter, I am concerned with something closer to accountability blameworthiness than attributability blameworthiness. Put differently and more loosely, in this chapter, I am concerned with blame for one's action or omission in contrast to blame, or criticism, for one's character or care.

Related to this, rather than using the distinction between accountability and attributability as a point of departure, and explicating those notions, some theorists prefer to begin with two conditions which they believe are separately necessary and jointly sufficient for blameworthiness – a *control condition* and an *epistemic condition* – and to explicate those.⁷⁷ The former condition has to do with whether the agents possess control over their actions or omissions and the latter with whether they are aware of their actions or omissions and their consequences. It is unfitting to blame an agent (and it is unfair to blame someone overtly) for something she lacked control over and for an action or omission she did not know she was performing, is the idea here.

⁷⁷ See, for example, Clarke (2017) and Zimmerman (1997; 2022).

Although I will explicate both conditions in the sections to come, it could be said that my focus in this chapter is mainly on the epistemic condition and less so the control condition.

Before moving on, and to be clear, although it may sometimes be fair to say that defenders of the accounts that I will discuss below talk past each other because they are interested in when different kinds of blame response are fitting, in what follows, I assume that defenders of *both* the witting wrongdoing view and the wrong-making view are concerned with whether blame (as I understand blame) is fitting. So, as I have set things up, these defenders are not concerned with different conceptions of blame, or blameworthiness, and do not therefore talk past each other.

3.2 *De Dicto* and *De Re*

The distinction between *de re* and *de dicto* is famously associated with Michael Smith's work (1994). He touches on this distinction when he defends his view of the kind of motivation possessed by the virtuous agent. He writes:

Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, not just one thing: doing what they believe to be right, where this is read *de dicto* and not *de re*. Indeed, commonsense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue. (Smith 1994: 75)

Roughly, Smith argues that virtuous agents perform actions on the basis of the features which make them right (*de re*), not on the basis of what they believe or judge to be right (*de dicto*). To perform actions or omissions solely on the basis that they are right reveals something defective in one's character, is fetishist, according to him. This observation is similar to Bernard Williams' (1981) famous "one thought too many" objection: there is something morally unattractive about a husband who saves his wife because she is his wife *and* because (he thinks) it is right to save his wife – he should be motivated solely by the thought that she is his wife.

The distinction between *de re* and *de dicto* is complex. It is not my aim to provide a clear statement, or detailed examination, of it here. To me it is sufficient to point

out that, within the domain of blameworthiness, the distinction has been used mainly to mark the contrast between agents who are motivated by the feature(s) which makes their action or omission wrong (*de re*) and agents who are motivated by their belief, or judgment, that their action or omission is wrong (*de dicto*). The witting wrongdoing that I will discuss in §3.4 might be interpreted as endorsing some form of *de dicto* concern, as it states that agents are blameworthy for having performed actions or omissions if, and only if, they willingly performed those actions or omissions despite *consciously believing it was the wrong thing to do*.⁷⁸ The wrong-making view, which I will discuss in §3.5, can be interpreted as endorsing a *de re* concern, since it states that agents are blameworthy for having performed actions or omissions if, and only if, *they performed them deliberately and were motivated by the reasons which made their actions or omissions wrong*.

3.3 Two Cases

In Chapter 1 (§1.3), I said that I would take it as a given that an adequate account of blame should account for our intuitions. If a theory cannot account of our intuitions adequately, we have a *pro tanto* reason to reject it. Conversely, if a theory can account for our intuitions adequately, we have a *pro tanto* reason to favour it. Put differently, I will assume that an account of what makes agents blameworthy for particular acts or omissions which does not yield conclusions that align with our intuitions, and that cannot provide a good explanation for why our intuitions are misplaced, must be flawed in some way.

Below, I present two cases that any account of blameworthiness needs to get right. Although not all blame scholars will share my intuitions, I believe that many blame scholars (and others alike) will.

Essentially, the first case concerns an agent who engages in conduct that is clearly wrong, or bad, knowing full well the suffering she thereby causes and so on, but nonetheless believes, at the time of action or omission, that what she is doing is right, or good.

⁷⁸ Again, I mean *overall morally* wrong, not *prima facie* morally wrong.

Bullying: Andy recently started studying at a university. Before this, he had no friends. However, since he started studying he has gained several friends, or, more accurately, he is one of a group of people studying on the programme he is enrolled in. He is thrilled over finally having met some friends. The leader of the group is very charismatic. He convinces the others in the group to join him in bullying Dwight – another person enrolled at the same programme, but not one of their group – so that Dwight, in the end, will leave their program. Andy believes that his and his fellow group members’ bullying makes Dwight suffer significantly. Despite that, he believes that what he and his fellow group members are doing is right, or good, because, as the group’s leader has said, Dwight is “poison” to the programme, “boring”, makes the programme they are enrolled in worse.⁷⁹ Thus he bullies Dwight, motivated by the thought that he is ridding the programme he is enrolled in of something bad.

The case of Andy is not unique. It is easy to find similar cases both in real life and hypothetically. Think, for instance, of all the SS-Officers who knew full well the suffering their acts caused and so on, yet freely persisted with their atrocities with the motive that it was the right thing to do. Think also, as Zimmerman (2022: 23) has remarked, of all the members of ISIS who routinely engaged in atrocities knowing full well the suffering their acts caused and so on, yet freely continued in their terrible behaviour with the motive that it was the right thing to do (because it was in line with God’s will).

Common sense dictates that agents such as Andy are fitting targets of blame for engaging in cruelties even though they can truthfully claim that they did not, at the time of acting or omitting, believe that they did anything wrong. Indeed, they seem to be prime examples of blameworthy agents. Encounters with agents such as Andy would probably lead us to adopt the blame sentiment towards them on the basis of their appalling conduct, and we would not find it odd if someone, especially a victim, were to harbour the blame sentiment towards them on the basis of their behaviour.

The other case I wish to consider can be interpreted as the reverse of the first case. It concerns an agent who engages in what is clearly a good, or right, action, knowing full well the relevant right-making or good-making facts, yet at the time

⁷⁹ In what follows, by “right” I mean overall morally right.

of acting or omitting believes that she is doing wrong. Arpaly's Huckleberry Finn is of such an agent:

Huck Finn befriends Jim, a slave, and helps him escape from slavery. While Huck and Jim are together on a raft used in the escape, Huck is plagued by what he calls "conscience." He believes, as everyone in his society "knows," that helping a slave escape amounts to stealing, and stealing is wrong. He also believes that one should be helpful and loyal to one's friends, but loyalty to friends is outweighed by some things, such as property rights, and does Miss Watson, Jim's owner, not have property rights? Hoping against hope to find some excuse not to turn Jim in, Huck deliberates. Like many children (and adults), Huck is not very good at abstract deliberation, and it never occurs to him to doubt what his society considers common sense. Thus, he fails to find a loophole. "What has poor Miss Watson done to me," he berates himself, "that I can see her [slave] go away and say nothing at all?" Having thus deliberated, Huck resolves to turn Jim in, because it is "the right thing." But, along comes a perfect opportunity for him to turn Jim in, and he finds himself psychologically unable to do it. He accuses himself for being a weak-willed boy, who has not "the spunk of a rabbit" and cannot bring himself to do the right thing, and eventually shrugs and decides to remain a bad boy. (Arpaly 2002: 228)

Intuitively, it is not fitting for anyone to blame Huck for helping Jim escape. If anything, it seems fitting for anyone to praise him – and this is so even though he believed, at the time of acting, that helping Jim was wrong, or bad.⁸⁰ Encounters with agents such as Huck would probably not lead us to adopt the blame sentiment towards them on the basis of their deeds, and we would find it odd if anyone were to adopt the blame sentiment towards them on that basis, especially Jim.⁸¹

The case of Huck Finn is controversial. As it is a case drawn from a novel, there are various interpretations of Huck's psychology when helping Jim escape. On some, it turns out to be clear that Huck is not blameworthy. On others, he is clearly, or at least somewhat, blameworthy. So, we should not rely too much on it. Let us consider a less controversial case designed to perform the same role as the one involving Huck:

⁸⁰ In Appendix, I will briefly discuss whether Huck is praiseworthy for helping Jim escape.

⁸¹ This is compatible with saying that other responses can be fittingly directed upon Huck, such as (mild) criticism (since Huck is not very good at moral reasoning).

Refugee: Pam is on Samos filming a documentary about the refugee camps there. When she is not abroad working, she lives in a flat in a good neighbourhood in Stockholm, Sweden. During the filming, a ten-year-old orphan boy at the camp asks her if she could take him with her by hiding him in her car when she drives home to Stockholm and help him obtaining permanent residence in Sweden. Pam believes that is illegal to smuggle people to Sweden. In addition, she believes that it is wrong to act illegally. However, she also believes that she can easily smuggle the boy to Sweden, and that not to bring the boy with her would involve handing over him to a camp that is not at all safe for young orphan boys – the camp is targeted by people searching for young boys to recruit into human slavery, and such like. And she believes that the boy is just like any other ten-year-old – in need of care and safety. In the end, she decides to hide the boy in her car with the motive of helping him. However, while concealing the boy throughout the ride back home to Stockholm, she consciously believes that what she is doing is wrong.

Common sense dictates that Pam is not a fitting target of blame for smuggling the boy to Sweden. If anything, she, like Huck, seems to be praiseworthy, even though she did something which, at the moment of acting, she believed was wrong. Further, it is hard to see that we ever would acquire the blame sentiment as a result of encountering Pam and learning about her conduct, and we would find it odd if someone, especially the boy she helped, were to adopt the blame sentiment towards Pam on the basis of her having helped the boy move to Sweden.

In sum, we believe, roughly, that agents who do something bad, or wrong, but believe they are doing something good, or right, are fitting targets of blame, and that agents that do something good, or right, but believe they are doing something wrong, or bad, are not fitting targets of blame. The challenge for accounts of blameworthiness is to account for these intuitions adequately, or otherwise provide a good explanation of why we should disregard our intuitions in the cases above. As will we see in a moment, accounting adequately for the cases presented above is harder than we might, taking the facts at face value, expect.

3.4 The Witting Wrongdoing View

In this section, I present and evaluate the witting wrongdoing view, asking how well it accounts for the cases presented in the previous section. I rely mainly on the witting wrongdoing view defended and developed by Zimmerman in several places (e.g., 1988; 1997; 2008; 2022). Ishtiyaque Haji (1997) can also be interpreted as endorsing some version of the witting wrongdoing view. I believe that much of what I say about the explication below applies to his view as well.⁸²

The witting wrongdoing view imposes two necessary conditions: (i) the agent A must have exhibited control over her act or omission ϕ , and (ii) A must have had a belief with the content “ ϕ is wrong” when ϕ -ng.⁸³ Below, I say more about (i) and (ii), starting with (i).

By “control over her action ϕ ”, defenders of the witting wrongdoing view mean that the agent exhibited *volitional control* over her act or omission – i.e., the sort of control we exercise when we do, or omit to do, something deliberately, or intentionally. The act or omission is “up to me”. I can see to it that the act is performed or omitted, or I can see to it that it is not performed or omitted. And I can decide when to act or omit to do something (now or later). In contrast, we do not exhibit any voluntary control when we slip on a banana peel. Nor do we exhibit volitional control over attitudes, such as beliefs. They are commonly thought to be adopted “automatically” in response to the “right kinds of reason” – which is to say that while we do not have volitional control over our attitudes, we presumably have *attitudinal* or *rational control* over them.⁸⁴ For example, my belief that it is morning is an “automatic” response I have as result of waking up and seeing the daylight outside my window.⁸⁵

⁸² Some add Rosen (2004; 2008) to the list of defenders of the witting-wrongdoing view. I believe that Rosen’s view is significantly different from Zimmerman’s view. That is why I have not included Rosen to the list of defenders of the witting wrongdoing view.

⁸³ Again, here I mean *overall morally* wrong, and not *prima facie* wrong.

⁸⁴ For more on rational, or attitudinal, control, see McHugh (2017). In Chapter 4 (§4.1), I will say more about what rational control is.

⁸⁵ While some of those discussing the witting wrongdoing claim that in order to exhibit control over one’s action, one must also be aware of its deontic status, not all do. I will assume that agents need not exhibit such awareness. It is enough that they are aware that they have made a decision. For the first see Rudy-Hiller (2017), for the second, see Zimmerman (2022).

The inclusion, by defenders of the witting wrongdoing view, of a condition concerning control preserves the widely shared theoretical belief mentioned in §3.1: namely, that it is unfitting to subject an agent to blame for something she lacked control over.

Moving to (ii), this condition states that the agent must have had a belief with the content “ ϕ is wrong” when ϕ -ing. It does not demand that the belief is true or that she knows and is applying the correct fundamental moral theory. She can be blameworthy for having acted despite her false belief that ϕ is wrong as well as for having acted with the true belief that ϕ is wrong, for example. Defenders of this view think that there is something blameworthy about agents who do not respect their own beliefs about morality’s demands – even when those demands are false.

By “belief”, defenders mainly mean *conscious* or *occurrent* belief, not *dispositional* or *unconscious* belief. Zimmerman provides a good explication of “occurrent belief”. He says:

[...] occurrently believing a proposition p consists in being conscious of or, in the case of true propositions, aware of the truth of p . This requires attending to p to some degree, but the degree may be minimal; for that is all that considering (entertaining, advertent to, thinking about) p requires. (Zimmerman 2022: 117)

Roughly, the belief that ϕ is wrong must in some sense be before the agent’s mind when she is acting or omitting. Zimmerman notes, however, that sometimes agents need not hold the belief before their minds when acting or omitting in order to be blameworthy. He writes:

The one possible exception is this: it may be that routine or habitual actions are performed for reasons to which one does not advert. (Zimmerman 1997: 422)

That said, most defenders of the witting wrongdoing view, including Zimmerman, assume that the relevant belief about wrongness is conscious, and I will do so, too, for the rest of this thesis.

The reason why defenders of the witting wrongdoing view are concerned with conscious beliefs is that as long as our belief about the deontic status of our act or omission is in the “background”, merely dispositional, we cannot use it when we are deciding what to do or omit – i.e., we can decide to act or omit neither *on* the belief nor *despite* the belief. According to defenders of the witting wrongdoing view,

one incurs [blameworthiness] for one's action only if one's belief concerning wrongdoing plays a role in the reason for which one performs the action. (Zimmerman 1997, 421-422)

And a belief plays a role in one's decision, defenders of the witting wrongdoing view argue, when one decides to act or omit either on the basis of belief or despite the belief.

Now to the important clarificatory question: Why must the relevant belief have the content that the action or omission is wrong? To answer this question, we need, first of all, to note a few features of ignorance.

Ignorance is a lack of knowledge, belief or awareness about something. In the debate over ignorance and blame, it is common to distinguish between *factual ignorance* and *moral ignorance*. The former denotes lack of knowledge, belief or awareness relating to something empirical, or the right-making or wrong-making features of the action or omission, while the latter denotes lack of knowledge, belief or awareness about moral matters, such as lack of knowledge about what you morally ought, or ought not, to do in a given situation, or a failure to believe that your act or omission is wrong or right.

The distinction above suggests that moral ignorance is not at all dependent on factual ignorance. That is problematic, as factual ignorance usually explains one's moral ignorance. For example, the fact that I am not aware that my action will very probably make you suffer significantly explains why I do not believe that my action is wrong.

Now, according to defenders of the witting wrongdoing view, factual ignorance is not directly relevant for one having an excuse (I will say more about excuses in §3.6.2). Only moral ignorance is directly relevant for one having an excuse, and even one's moral ignorance will not excuse one unless one is to blame for it. To illustrate, consider Andy in §3.3. Suppose he is ignorant of the fact that his action will probably make Dwight suffer significantly but nonetheless believes that bullying Dwight is wrong. According to defenders of the witting wrongdoing view, he will then not be excused for bullying Dwight as a result of his factual ignorance. Rather, he will be blameworthy (granting that he also bullied Dwight voluntarily). This result is intuitive.

In contrast, let us consider the case of Andy as it was presented in §3.3. In that presentation, Andy is not factually ignorant. Rather, he is morally ignorant. That is, Andy knows that his action will probably make Dwight suffer significantly, but he fails to believe that it is therefore wrong. According to defenders of the witting wrongdoing view, Andy can be excused for bullying Dwight, since he fails to believe that it is wrong for him to bully him. At least, Andy is blameless unless we can establish that he is blameworthy for failing to believe that it is wrong to bully Dwight. If it turns out that Andy is blameworthy for failing to believe that bullying Dwight is wrong, he might be *indirectly* blameworthy for bullying Dwight.

Crudely put, defenders of the witting wrongdoing view hold that Andy is indirectly blameworthy for his failing to believe that bullying Dwight is wrong when that failure is a result of some prior blameworthy action or omission he performed, and that he also at that time (in some sense) foresaw would have the consequence that he will have false moral beliefs about it being right, not wrong, for him to bully Dwight. And an agent *A* is blameworthy for some action or omission if, and only if, (i) *A* had voluntary control over the action or omission, and (ii) *A* performed the act or omission despite consciously believing it to be wrong to do so. As Zimmerman puts it:

[...] *all* blameworthiness rests on, or is rooted in, *non-ignorant*, that is, *witting* [and willing] wrongdoing. (Zimmerman 2022: 21)

So, in order for Andy to be indirectly blameworthy for bullying Dwight as a result of moral ignorance, his failure to believe that bullying Dwight is wrong must be a foreseeable result of some prior act or omission of his, and it must also be the case that, at the earlier time, he acted or omitted voluntarily despite consciously believing that it was wrong to act or omit in that way. As defenders of this view predict, it is doubtful whether we will succeed in finding such an action or omission in the history of agents like Andy. Probably, we will not. Therefore, it is likely that fewer people than we are inclined to think are blameworthy are in fact blameworthy.

Given this, we can summarise the witting wrongdoing view as follows:

The witting wrongdoing view: an agent *A* is blameworthy for having performed a particular action or omission ϕ if, and only if, (i) *A* had

volitional control over φ , and (ii) A φ -ed despite consciously believing it to be wrong to φ .

I now turn to investigate whether we can rely on the witting wrongdoing view to make adequate sense of the cases in §3.3.

As is obvious from the above, the witting wrongdoing view entails that agents like Andy are not blameworthy for the atrocities or bullying they engage in. This is because they lack the relevant conscious belief at the moment of acting – i.e., the belief that it is wrong to engage in the atrocities or bullying. Thus, the witting wrongdoing view fails to classify agents we intuitively believe are blameworthy as blameworthy.

One might argue that the above conclusion is too quick. Although Andy is not *directly* blameworthy for bullying Dwight, he could be *indirectly* blameworthy for failing to believe that bullying Dwight would be wrong. To establish this claim, as will be recalled from the comments above, we need to find an act or omission in Andy's history that he foresaw (in some sense) would have the consequence that he has false beliefs that it was right to bully Dwight, and he must additionally be blameworthy for that act or omission – i.e., he must have performed that act or omission voluntarily, despite consciously believing that it was wrong for him to do so.

The case, as presented in §3.3, is silent on whether Andy performed any such act or omission prior to bullying Dwight. For example, we do not know whether Andy decided to trust the leader's opinion of Dwight even though he foresaw that it might result in him having false beliefs about it being right, not wrong, for him to bully Dwight, and even though he believed at that time that it is wrong to, say, trust charismatic leaders. That said, I believe our intuitions about Andy stand – regardless of whether we succeed in finding such an event in his history. Even if it turns out that he did not perform the relevant act or omission, I believe we still would find it fitting to blame Andy for bullying Dwight. And we would not find it odd if anyone, including Dwight, were to blame Andy for his act of bullying despite him not being indirectly blameworthy (on the witting wrongdoing view) for his moral ignorance.⁸⁶

⁸⁶ In §3.6.2 I provide a brief explanation of why it is fitting to blame Andy.

Another reply would be to argue that although Andy does not, at the time of acting, *consciously* believe that what he is doing is wrong, he believes it *unconsciously*.

As the case is presented in §3.3, it is also silent on whether Andy harbours any unconscious beliefs with the content “engaging in the bullying is wrong”. If he does, and if we modify the witting wrongdoing view to state that we can also be blameworthy for acting in contravention of our unconscious beliefs with the content that “ ϕ -ing is wrong”, then we can account for our intuition about Andy. However, I believe that this strategy is unhelpful. We can stipulate that Andy does not harbour that kind of unconscious belief. If we do, I think we will still believe that it is fitting for anyone to blame Andy for bullying Dwight, and we cannot rely on the witting wrongdoing view to reach that conclusion.⁸⁷

The witting wrongdoing view also fails to account for the second case presented in §3.3 concerning agents who voluntarily do good, or right, while consciously believing that they are doing wrong. More precisely, the view entails that Huck (depending on one’s interpretation of the case) and Pam are blameworthy for helping, respectively, Jim and the boy, since they both voluntarily help them while at the time of action consciously believing that it was wrong to do so.

In sum, acceptance of the witting wrongdoing view obliges us to give up some of our intuitions about blameworthy and non-blameworthy conduct. We need to radically revise our inclination to blame agents such as Andy and not blame agents such as Pam. In addition, since it is plausible to suppose that few agents voluntarily do what they consciously believe is wrong, we would probably have to accept the claim that not as many agents we believe are blameworthy are actually blameworthy.

Defenders of the witting wrongdoing view are aware that their account has unintuitive and revisionary implications. That said, they welcome the results – that Huck and Pam are blameworthy, and that Andy and other agents like him are not, on their view. As regards Huck and Pam, they invite us to focus on the

⁸⁷ Another response is to claim that agents like Andy represent exceptions to Zimmerman’s restriction to occurrent beliefs presented above. More precisely, such agents might be blameworthy because they engage in deliberate wrongdoing in a routine, or habitual, and hence inadvertent, manner. However, Zimmerman does not appear to think that agents like Andy are exceptions, he acknowledges that his verdict that agents like Andy are not blameworthy can be viewed as “repugnant” (Zimmerman 2022: 23).

fact that Huck and Pam believe that they do something wrong, and not the fact that they acted rightly, or did something good, and that Andy acted wrongly or badly. As stated above, acting despite one's beliefs about moral demands signals some kind of disrespect for that authority – a blameworthy quality of mind – even when our beliefs about morality's demands are false.

The cases of Huck and Pam in §3.3 are narratives in which agents do not respect their own beliefs about what morality demands of them, but where we nonetheless do not think that the agents are blameworthy *at all* for their actions. So, this reply is not so convincing, or, at least, there is a clash of intuitions.⁸⁸

Defenders of the witting wrongdoing view tend to point out the fact that Huck and Pam are fitting targets of blame, and that Andy is not, does not entail no other fitting responses can be fittingly adopted towards them, or that we must let agents like Andy “off the hook”. It may be fitting, as Zimmerman (2008; 2022) holds, at the same time to admire Huck for his bravery, or to feel disgusted by, or contempt for, Andy given his bad character traits.

Although negative responses other than blame can be fittingly directed upon Andy, and although positive responses differing from blame can be fittingly directed upon Pam and Huck, I believe our intuition is that it is not fitting *at all* to blame Pam and Huck, and that it is fitting to blame Andy. That other responses also are fitting does not help to dislodge this intuition; it is beside the point that other responses also are fitting.

The final reply I wish to consider is to question the theoretical claim, presented in §3.3 and Chapter 1 (§1.3), that an adequate theory of what makes agents blameworthy must be in line with our intuitions. One might argue that it should be the other way around: our accounts of blameworthy conduct should inform our practice of blame, not the other way around. On this approach, our practice of blame excuses agents insufficiently, because it is at odds with the witting wrongdoing view.

⁸⁸ An idea I would like to have explored, given more space, is the claim that Huck and Pam are *pro tanto* blameworthy, or *locally* blameworthy, but not *overall* blameworthy, or *globally* blameworthy. There might be a moral principle stating that it is wrong to do what one believes is wrong. Given that Huck and Pam violated that principle by acting as they did, we have a *pro tanto* reason to blame them for their acts. However, that *pro tanto* reason is not strong enough to make them overall blameworthy for their acts.

Although this reply is possible and interesting, in this thesis, I assume that theoretical claim I presented in Chapter 1 and restated in §3.3 is true, and I do not wish to question it here – that is a task for future research.

To conclude, although defenders of the witting wrongdoing view can offer various replies to the claim their view entails unintuitive verdicts on the cases presented in §3.3, I believe that many of us would take the replies to be less than fully persuasive. There is something wrong with an account of blameworthiness that entails that Pam, but not Andy, is to blame. This does not mean I reject the witting wrongdoing view. My view is merely that I will not favour it until it is shown that there is no other account of what makes an agent blameworthy for having performed a particular action or omission that makes better sense of the cases provided in §3.3 and is otherwise plausible. Next, I consider whether the wrong-making view makes better sense of the cases in §3.3.

3.5 The Wrong-Making View

In this section, I present and evaluate the wrong-making view, asking how well it accounts for the cases in §3.3. As with the witting wrongdoing view, there are different formulations of the wrong-making view.⁸⁹ Although Nomy Arpaly and Timothy Schroeder (2013) are the principal defenders of this view, my starting point is Julia Markovits' (2012) formulation of it.⁹⁰ That said, I believe that most of what I say below can be applied to the views of Arpaly and Schroeder, and to the views of other defenders of the wrong-making view as well. Put differently, it could be said that my aim in this section, and those that follow, is to formulate and defend a novel wrong-making view.

Here is a first sketch of the wrong-making view:

The wrong-making view: an agent *A* is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) *A* had volitional control

⁸⁹ Compare, for instance, Markovits (2012), Arpaly (2002), Arpaly and Schroeder (2013) and Alvarez and Littlejohn (2017).

⁹⁰ It should be mentioned that Markovits focuses on praiseworthiness. However, she has formulated a view of blameworthiness, and that is my main point of departure. See Appendix for her view of praiseworthiness.

over $\varnothing$, and (ii) A 's motivating reason(s) for $\varnothing$ -ing coincides with the normative reason(s) that makes it wrong for A to $\varnothing$.

The wrong-making view clearly needs some unpacking. Below, I briefly explain what I mean by "volitional control", "normative reason(s)", "motivating reason(s)", and "wrong".

By "volitional control" I mean, as the reader might recall from §3.4, the sort of control we exercise when we do something willingly or intentionally. I can decide when to perform the action or omission (now or later). The action is "up to me": I can see to it that it is, or is not, performed.⁹¹

I take "normative reason(s)" to be a distinct fact, or set of facts, that stand in a certain relation to an agent and a response. For example, the fact that it is raining outside is a reason for me to bring an umbrella with me when I go out.

By "motivating reason(s)" I mean a reason of the kind we indicate when asked "What was her reason for acting as she did?"⁹² This reason needs not be provided by a distinct fact, or set of facts. In contrast, normative reasons are provided by facts. For example, an agent can engage in atrocities for the reason that she is thereby acting in accordance with a certain zombie's will. But since there are no zombies and hence no zombie's will, this motivating reason is not a normative reason to engage in the atrocities.⁹³ Derek Parfit says something similar here:

If you ran away from the angry snake, your motivating reason would be provided by your false belief that this act would save your life. But, as I have said, you have no normative reason to run away [because the snake in fact only attacks moving

⁹¹ Some defenders of the wrong-making view, for example, Arpaly and Schroeder (2013), do not explicitly include a control condition. However, I believe it is a virtue of the account that it explicitly preserves the theoretical belief that it is unfitting to blame an agent for an act or omission that she did not have control over.

⁹² Does it matter how many normative reasons, or wrong-making features, the agent is motivated by? Arpaly and Schroeder (2013) distinguish between agents motivated by *all* wrong-making features and agents motivated by *some* wrong-making features. They say the first type of agent exhibits *complete ill will* and the second agent *partial ill will* (in a moment, I will say more about the way in which Arpaly and Schroeder understand "quality of will"). Both kinds of agent, on their view, are blameworthy, but presumably to different degrees. Although an interesting question, I will not say more about it here since it has no bearing on whether the wrong-making view can handle the cases presented in §3.3.

⁹³ Perhaps it is merely an *apparent* reason to engage in the atrocities. For more about apparent reasons, see Parfit (2011).

targets]. You merely think you do. In an example of a different kind, we might claim: ‘His reason was to get revenge, but that was no reason to do what he did’. (Parfit 2011: 37)

It should be emphasised that motivating reasons are not the same as *explanatory reasons*. The latter are the reasons that explain why the agent acted as she did. For example, my lack of good sleep explains why I uttered a rude remark to you.⁹⁴

Moving to “wrong” – and, more precisely, to the notion that the agent’s motivating reason(s) for ϕ -ing “coincides with the normative reason(s) that makes it wrong for A to ϕ ” – there is a lively debate as to whether or not rightness and wrongness depend on an agent’s epistemic circumstances.⁹⁵ *Objectivists* about rightness and wrongness believe that rightness and wrongness do not depend at all upon which facts are epistemically available to agents. Rather, rightness and wrongness depend on all the facts.⁹⁶ If it is wrong for an agent A to perform an action or omission ϕ , this is because ϕ has some property P which renders it so, irrespective of whether A knows, or is able to know, that ϕ has P .

In contrast, *subjectivists* about rightness and wrongness believe that rightness and wrongness do depend on which facts are epistemically available to the agent.⁹⁷ Subjectivism comes in different forms. Some subjectivists believe that it is the agents’ doxastic states that determine what it is right or wrong for them to do. Others – specifically, *perspectivists* – think, that it is facts the agents are aware of, or their evidence, that determine what it is right or wrong for them to do. In short, by “aware” I mean that the agent is either conscious of the fact or the fact is self-accessible to her – i.e., she can know, recognise or tell the fact solely by means of reflection and self-examination. For example, I am aware of the fact that it is your birthday today either if I am conscious of it or I can easily come to recognise that it is your birthday today by means of reflection.⁹⁸

⁹⁴ For a good summary of these kinds of reasons, see Werkmäster (2019).

⁹⁵ Here, again, I mean *overall moral* right or wrong, not *prima facie* right or wrong.

⁹⁶ See, for example, Thomson (1990).

⁹⁷ See, in particular, Kiesewetter (2017), Scanlon (2008), and Skorupski (2010a).

⁹⁸ One might complicate the picture further by adding *prospectivism* to the list. For a defender of prospectivism about rightness and wrongness, see Zimmerman (2008). For reasons mainly having to do with time and relevance, I will not consider prospectivism about rightness and wrongness here.

It lies outside the scope of this thesis to provide an adequate defence of either of these views (if I should provide one or instead just be a pluralist about rightness and wrongness – i.e., accept that there are various notions of wrongness and rightness in play).⁹⁹ Nevertheless, I believe objectivism is open to a serious challenge, and that this challenge speaks forcefully in favour of some sort of subjectivism. The challenge is based on a famous case provided by Frank Jackson (1991):

Drugs: Jill is a physician who has to decide on the correct treatment for her patient, John, who has a minor but not trivial skin complaint. She has three drugs to choose from: drug *A*, drug *B*, and drug *C*. Careful consideration of the literature has led her to the following opinions. Drug *A* is very likely to relieve the condition but will not completely cure it. One of the drugs *B* and *C* will completely cure the skin condition; the other though will kill the patient, and there is no way that she can tell which of the two is the perfect cure and which is the killer drug. (Jackson 1991: 462)

It is clearly right for the doctor to prescribe drug *A*, and this despite the fact that she knows that it is not the “objectively right” or best option. For her to prescribe either of the other two drugs would be reckless, and thus wrong. Prescribing either of drugs *B* or *C* would be reckless because her available information does not allow her to discriminate between them. For all she knows, she could be prescribing poison, and curing a minor skin condition is not worth the risk of death. We can call this challenge the *recklessness objection*.

I believe the recklessness objection nicely illustrates why rightness and wrongness depend on which facts are epistemically available to agents. In what follows, I assume that some form of subjectivism about rightness and wrongness is correct. I do not have the space to expound an argument showing which kind of

⁹⁹ A reason not to be a pluralist arises from cases of deliberation where the objective and subjective views point in different directions, and we want to know which direction is more important. Afterall, I cannot both ϕ and not- ϕ . As Kieseewetter argues: “The point of deliberation is, after all, to guide rational decision-making and belief-formation, and it can fulfil this function only if there is one univocal sense of ‘ought’ that figures in deliberative conclusions rather than a variety of potentially conflicting sense. It is perfectly consistent to believe “I ought to ϕ , relative to *X*” and “I ought to ϕ , relative to *Y*”, but one cannot rationally intend both to ϕ and not to ϕ . There must be *one* sense of ‘ought’, the belief in which is the relevant one for deliberative conclusions. There is thus a substantial question of whether the deliberative ‘ought’ is sensitive to the epistemic circumstances of the agent or not – a question that cannot be divided by distinguishing different senses of ‘ought’” (Kieseewetter 2017: 197).

subjectivist we should be. However, in what follows, I will develop my argument on a perspectivist basis.

So, it will be facts agents are aware of that decide whether or not the performance of a particular action or omission is wrong. For example, and crudely put, the fact that ϕ -ing is expected to cause unnecessary suffering, which I am aware of, gives me a normative reason to not- ϕ . Given that there are no other normative reasons speaking for or against ϕ -ing, this normative reason also amounts to it being wrong for me to ϕ .¹⁰⁰

Before moving on, I need to stress that the distinction between rightness and wrongness should not be conflated with the distinction between *goodness* and *badness*, or between what is *desirable* and what is *undesirable*. Eduardo Rivera López helpfully illustrates the difference here:

It is undesirable that I fail to save all the starving people in the world, but my failure to do so is not wrong (in this case, because it would be impossible for me to avoid). Judgments about what is good and bad (or desirable and undesirable) need not have any strong link to human limitations and capacities. But judgments about what is right and wrong should have at least some connection with them. (Rivera López 2006: 139-140)

In other words, it is, for example, possible that a right action or omission is undesirable, and a wrong action or omission desirable.

Before asking how the wrong-making view handles the cases presented in §3.3, it will be helpful to compare it with the witting wrongdoing view. This will make both accounts clearer and the differences between them more apparent.

Unlike the witting wrongdoing view, the wrong-making view requires blameworthy agents to act *wrongly* or to omit *wrongly* in virtue of (ii), which states that *A*'s motivating reason(s) for ϕ -ing coincides with the normative reason(s) that *makes it wrong* for *A* to ϕ . Defenders of the witting-wrongdoing view do not demand that: they demand that blameworthy agents *believe* they act or omit wrongly. Further, they think we can be blameworthy not just for actions or

¹⁰⁰ It should be stressed that facts such as " ϕ is wrong" do not count as normative reasons against performing ϕ on this view. Rather, only epistemically available facts in virtue of which one's action is wrong count as normative reasons against performing ϕ .

omissions that are wrong, but also for actions or omissions that are right or neutral. In section §3.6.1, I will say more about the relation between blameworthiness and wrongness, and the possibility of blameworthy right doing.

Although both accounts agree that blameworthy agents must exhibit volitional control over their actions and omissions, they differ in the way they understand the epistemic condition: while defenders of the witting wrongdoing view posit as a necessary condition that blameworthy agents believe that their acts or omissions are wrong, defenders of the wrong-making view do not set such a condition. Instead, the necessary condition they posit is that blameworthy agents are (aware and) motivated by the normative reasons that make their actions or omissions wrong.

Another difference is that while defenders of the witting wrongdoing view hold that moral ignorance excuses the agent (unless she is to blame for her moral ignorance), defenders of the wrong-making view deny this.

Relatedly, defenders of the witting wrongdoing view hold that agents cannot be directly blameworthy for acts done out of moral ignorance – they can only be indirectly blameworthy for acts done out of moral ignorance. But defenders of the wrong-making view hold that agents can be directly blameworthy in that way. In §3.6.2, I will say more about the wrong-making view and excuses, and why defenders of the wrong-making view believe that it is fitting to blame agents for morally ignorant conduct.

I now turn to consider how the wrong-making view handles the cases in §3.3.

Compared with the witting wrongdoing view, it delivers an intuitive verdict in the case involving Pam. Pam's action is not wrong. It is right.¹⁰¹ The facts of which she was aware of provide her with normative reasons that, together, make it is right, not wrong, for her to help the boy flee to Sweden. In addition, her motivating reason is to help the boy flee to Sweden. Thus, the wrong-making view trivially arrives at the right result with Pam – namely, that she is not blameworthy for having helped the boy flee to Sweden.

How about Andy? Andy is doing something wrong by bullying Dwight. The facts of which he is aware of, such as the fact that his action will very probably make

¹⁰¹ Or, perhaps, supererogatory. In what follows, I assume that her act is right. However, nothing important hinges on whether we should describe her act as right or supererogatory. The important claim is that her act is not wrong.

Dwight suffer greatly, provide him with normative reasons that, together, make it wrong for him to bully Dwight. However, he cannot be said to be motivated to bully Dwight by the belief that his act will cause significant suffering to Dwight. If we were to ask him, or any observer, what his motivating reason was, we would receive a reply referring to his false belief that he was ridding the programme he is enrolled in of something bad. Hence, the wrong-making view incorrectly implies that Andy is not blameworthy for engaging in the bullying, as his motivating reason does not coincide with the normative reasons that make it wrong for him to engage in the relevant behaviour. Like the witting wrong-doing view, the wrong-making view seems to imply that some of our ordinary judgments about blameworthy agents are false.

One might reply that a complete specification of Andy's motivating reason would reveal that he is in fact motivated by the aim to make Dwight suffer. More precisely, while it is true that Andy is not motivated to cause Dwight suffering *directly* or *non-derivatively*, it may well be true that he is motivated to cause Dwight suffering *indirectly* or *derivatively*. For example, he is motivated to *bring about Dwight's suffering as a means of ridding the programme he is enrolled in of something bad*.

There are several possible ways in which an agent could have been motivated by normative reasons, including derivatively or non-derivatively. However, I believe that the main point of contention has been that what is relevant is that the agent is non-derivatively motivated by the normative reasons.

Apart from the fact that being motivated to "cause Dwight suffering as a means to ridding the programme he is enrolled in of something bad" sounds like an unnatural and excessively complex motivation, it seems to me that Andy's being derivatively motivated by the wrong-making feature does not overlap perfectly with the normative reasons that make it wrong for him to engage in the bullying. What makes it wrong for him to engage in the bullying is chiefly the suffering he can be expected to cause to Dwight, not suffering *qua* instrument with which to rid the programme he is enrolled in of something bad.

Another reply, drawing inspiration from Arpaly and Schroeder (2013), is to claim that the wrong-making view, as stated above, is clearly inadequate. Agents are blameworthy not just when they are motivated by normative reasons that, together, make the actions they are performing, or the actions they omit from

performing, wrong, but also when they act or omit out of a lack of complete, or partial, good will – i.e., when they act or omit out of *moral indifference*.

There are many details in Arpaly and Schroeder's account to which I cannot do justice to here. Roughly speaking, they hold that an agent acts or omits out of moral indifference when her action or omission is rationalised by her lack of desire for the complete or partial right, correctly conceptualised. What is right and wrong is determined by the correct moral theory, according to Arpaly and Schroeder. For example, if utilitarianism is the correct moral theory, an agent who lacks a desire to maximise happiness lacks a desire for the right, correctly conceptualised.

It should be mentioned at this point that Arpaly and Schroeder explicate the wrong-making view in terms of their ideas about correctly conceptualised (intrinsic) desires rather than in the way I have done in this chapter. In short, for them, an agent *A* is blameworthy for having performed a wrong action or omission ϕ if, and only if, *A*'s ϕ -ing is rationalised by *A*'s ill will or partial ill will, or moral indifference. By "ill will" and "partial ill will" they mean having some or merely correctly conceptualised (intrinsic) desires for what is bad or wrong – e.g., desires to cause unnecessary suffering or to disrespect people.

With this in mind, we can modify the wrong-making view as follows:

*The wrong-making view**: an agent *A* is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) *A* had volitional control over ϕ , and (ii) *A*'s motivating reason(s) for ϕ -ing coincides with the normative reason(s) that makes it wrong for *A* to ϕ or *A*'s ϕ -ing is rationalised by *A*'s moral indifference.

The wrong-making view* can account for Andy. He seems to lack a desire for the complete or partial right, correctly conceptualised: thus, he does not have the desire not to engage in the bullying of Dwight because doing so involves not treating Dwight as an end in himself (assuming that Kantianism is the true moral theory). However, I do not think we should accept this modification.

As Errol Lord (2017) and Nathan Robert Howard (forthcoming) have argued, Arpaly and Schroeder's claim that the agent's desires deploy the concepts

belonging to the true moral theory makes their theory excessively demanding.¹⁰² It is doubtful whether anyone outside a philosophy classroom has desires that include concepts figuring in a correct moral theory, such as “end in themselves” or “maximise utility”. So, although we can handle the case of Andy by following Arpaly’s and Schroeder’s reply, we can only do so at the cost of making an implausible claim. It would be preferable to handle the case of Andy without having to rely on such a claim.

What seems to create the problems above is the fact that sometimes blameworthy agents act from motives that are not clearly bad – as, for example, Andy does. The modification of the wrong-making view I will discuss below begins with the idea that what is relevant for blameworthiness is not agents’ motives, but rather whether, at the time of acting or omitting, the agent is aware of the facts providing normative reasons that, together, make her action or omission wrong.

*The wrong-making view***: an agent A is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) A ϕ -ed despite, at the time of action or omission, being aware of the fact(s) that constitutes a normative reason which, taken together, makes it wrong for A to ϕ .¹⁰³

By “aware” in (ii) I mean, as noted above, that the agent is either conscious of the fact or the fact is self-accessible to her – i.e., she can know, recognise or tell the fact solely by means of reflection and self-examination. For example, I am aware of the fact that it is your birthday today either if I am conscious of it or I can easily come to recognise that it is your birthday today by means of reflection.

It is clear that Andy is blameworthy on the wrong-making view**. He acted wrongly by bullying Dwight, as will be recalled from the discussion above, since the facts he was aware of constituted normative reasons which, taken together, made it wrong for him to bully Dwight. And he willingly engaged in the bullying despite being aware of those facts.

¹⁰² Lord (2017) and Howard (forthcoming) focus on Arpaly’s and Schroeder’s view of praiseworthiness, but their complaint can readily be translated (as I have done here) to the similar view of blameworthiness taken by Arpaly and Schroeder.

¹⁰³ One can perhaps make the wrong-making view** more elegant, or simple, by rephrasing (ii) to (ii)*it is wrong for A to ϕ , or something like that.

Turning to Pam, the wrong-making view** can accommodate that case as well. At the time of acting, Pam is aware of the facts (including the fact that the boy is in need of help) that constitute normative reasons which, taken together, made it right and not wrong for her to help the boy flee to Sweden. And she deliberately helped the boy to move to Sweden.

The wrong-making view** can account for the cases presented in §3.3. So, at least in this respect, it is preferable to the witting-wrongdoing view and the wrong-making view. In addition, it does not depend on any implausible claims, such as the claim that agents need to have correctly conceptualised desires of what is in fact right or wrong, as the wrong-making view* does. Tentatively, I therefore conclude that the wrong-making view** is true.

Before concluding this chapter, I want to *briefly* consider how the wrong-making view** handles some other issues and claims in the blameworthiness debate.¹⁰⁴ Some of these issues have not been adequately discussed by defenders of the wrong-making view. If it can handle the challenges and claims, then we have additional support for the wrong-making view** – i.e., support beyond the fact that it handles the cases in §3.3.

3.6 Challenges

In this section, I consider three challenges to the wrong-making view**. These include the following: (i) we can be blameworthy for right, or neutral, actions or omissions, not just wrong actions or omissions (§3.6.1); (ii) the wrong-making view** and the standard view of excuses are in tension (§3.6.2); and (iii) it is unclear how we can rely on the wrong-making view** to make sense of the claim that an agent can be more or less blameworthy for an act or omission.

3.6.1 Blameworthiness and Wrongdoing

Some blame scholars, such as Zimmerman (2022) and Lord (2017), claim that blameworthiness and wrongness are two distinct concepts that should not be

¹⁰⁴ As we will see, the issues I will explore below are large and complex. Therefore, I can only offer a brief investigation of them here.

merged together. Whether an action is wrong or right is a *deontic* question, and whether an agent is a fitting target of blame for having performed it is a *hypological* question, on this view. Generally, the first question is determined by facts and the second by the agent's state of the mind when acting or omitting. Further, these scholars think that agents can be blameworthy for actions or omissions that are right or neutral, and not just those that are wrong.¹⁰⁵ Zimmerman puts these points as follows:

Whether an *action* is morally *right or wrong* is a deontic question; whether an *agent* is *worthy of praise or blame* is a hypological question. [...] I submit that it is simply a distraction, one that invites confusion, to concern oneself with the rightness of someone's action when trying to figure out whether that person is to be praised for that action. It is the agent's state of mind that matters in this regard, irrespective of the behaviour that results from or accompanies it. [...] I deny that one can only be [...] blameworthy for doing what is wrong. On the contrary, it is perfectly possible to be [...] blameworthy for doing what is right. (Zimmerman 2022, 321-322, emphasis original)

The wrong-making view**, just like other versions of the wrong-making view, requires that blameworthy agents act or omit wrongly in virtue of (ii), which states that *A* φ -ed despite, at the time of acting or omitting, being aware of the fact(s) that constitutes a normative reason(s) which, taken together, *makes it wrong for A to φ* . Hence, the worry is that the wrong-making view** does not respect the theoretical position above.

Note that this is not a worry for defenders of the witting wrongdoing view. They only demand that blameworthy agents *believe* they act or omit wrongly, not that they *in fact* act or omit wrongly.

I believe that the challenge above is serious. There do seem to be cases of agents who are blameworthy for having performed a right or neutral action. Consider, for example, the following variant of the classical trolley problem provided by Peter A. Graham:

¹⁰⁵ Interestingly, Mason (2019) has argued that subjectivism about wrongness correlates with blameworthiness, although she has a view of subjectivism which differs from mine. According to her, subjectivism concerns trying to do well by what she calls Morality and one is blameworthy when one fails to try to do well by Morality.

TROLLEY: James can divert an out-of-control trolley away from running over and killing five trapped track workers. If he does so, though, the trolley will run over and thereby kill Chris. James understands the situation and correctly concludes that it is morally permissible to divert the trolley. He does so, however, not to save the five, but to kill Chris, his despised uncle from whose death he will benefit. (Graham 2014: 395)

It is right, not wrong, of James's to divert the trolley, but he is blameworthy for doing so.

Consider also Skorupski's version of the drug case discussed in §3.5:

[...] suppose you administer a drug to a patient, and given your epistemic state [i.e., facts you are aware of], there is warrant for you to think it is a cure. But in fact you don't believe that; you *unwarrantedly* believe it to be a poison. You took yourself to be poisoning the patient; in fact, you cured him. (Skorupski 2010b: 167, emphasis original)

In the above case, which we can call POISON, it is right, not wrong, of you to administer the drug to the patient, but you are blameworthy for doing so. So, we have another case where blameworthiness and wrongness come apart.¹⁰⁶

There are different replies defenders of the wrong-making view** can provide to the cases above. One is to endorse the conclusion that James is not blameworthy for diverting the trolley, and you are not blameworthy for administering the drug, but argue that he and you instead have bad characters which render certain responses close to blame fitting, such as hate or (non-blaming) disgust. So, TROLLEY and POISON do not show that the wrong-making view** is false. Rather, they show that we need to be careful with regards to which kind of evaluation we are concerned with.

Another reply (given that the above reply is not convincing) is to show that James and you are blameworthy in TROLLEY and POISON. However, not for diverting the trolley or administering the drug, but rather for something else, such

¹⁰⁶ These cases also put pressure on the idea that blame, in its nature, is associated with the thought that the blameworthy agent *ought to have acted differently* or that the blameworthy agent acted wrongly or badly. For example, as the reader might recall from Chapter 2 (§2.1.5), Sher holds that to blame someone is to believe that she acted wrongly, or badly, and to desire that she acted differently.

as *trying to kill Chris* or *trying to poison your patient*. James and you are presumably aware of facts that provide normative reasons which, taken together, make it wrong for James to try to kill Chris, and wrong for you to try to poison your patient. Given that James tries to kill Chris, and given that you try to poison your patient, despite at the time of acting being aware of those facts, he and you are blameworthy according to wrong-making view^{**}. So, defenders of the wrong-making view^{**} can preserve the intuition that James and you are blameworthy for something in TROLLEY and POISON without accepting the claim that agents can be blameworthy for performing right or neutral actions.

The above reply is related to several difficult issues. Below, I present two.

In TROLLEY, the following descriptions of what James is doing all seem apt: (i) James is killing Chris; (ii) James is trying to kill his Chris; (iii) James is diverting the trolley; and (iv) James is trying to divert the trolley. Intuitively, (i)-(iv) describe the same action. Therefore, the actions described in (i)-(iv) should have the same deontic status. For example, it would be odd if it was wrong for James to try to kill Chris but not wrong for James to divert the trolley. Thus, there is a question about how we should distinguish between different actions and whether it is feasible to hold that what James is doing in (ii) is wrong but not what he is doing in (iii).

Second, it is unclear what “trying” denotes. It could refer to some *mental action* – a “setting oneself” or to forming an intention – or to some *doing* – e.g., searching for an opportunity to reach your goal, or preparing yourself to reach the goal.¹⁰⁷

To the point: for this reply to be convincing, these issues need to be addressed. I do not have the space to address these issues, nor do I have the time to evaluate whether the first reply to the challenge from blameworthy right doing is convincing. What we should take with us from this section is that the challenge that we can be blameworthy for right actions or omissions is a serious one for defenders of any wrong-making view.

¹⁰⁷ For more about the concept “trying”, see Werkmäster (2019) and Mason (2019).

3.6.2 The Wrong-Making View** and Excuses

Standardly, excuses leave wrongness in place but entail that the agent is not blameworthy for having performed the wrong action or omission.

Standard view of excuses: An agent A is excused for performing a certain action or omission ϕ if, and only if, (i) it was wrong for A to ϕ , and (ii) A is not at all blameworthy for having ϕ -ed.¹⁰⁸

According to the wrong-making view**, an agent A is blameworthy for a certain action or omission ϕ if, and only if, (i) A had volitional control over ϕ and (ii) A ϕ -ed despite being aware of the fact(s) that constitutes a normative reason which, taken together, makes it wrong for her to ϕ . Consider the revised case of Andy again (§3.4). Suppose that he was not aware of the fact that his action can be expected to make Dwight suffer significantly. If he indeed was unaware of that fact (and related facts), then presumably he did not act wrongly by engaging in the bullying. And he could not bully Dwight despite being aware of that fact. So, it is hard to see how, on this view, it could ever be the case that an agent has performed a wrong action, or omitted wrongly, but is not blameworthy for doing so. Put differently, Andy's factual ignorance "excuses" both by showing that he is not blameworthy for having bullied Dwight *and* by showing that his action was not wrong. The wrong-making view** thus puts pressure on the common claim that actions for which one is fully excused are nonetheless wrong.¹⁰⁹

There is a question about which account of wrongness the *standard view of excuses* assumes. As will be recalled from §3.5, there are various accounts of wrongness and rightness. If "wrong" in *standard view of excuses* means "objectively wrong", the wrong-making view** will not be in tension with the *standard view of excuses*. However, if "wrong" refers to "subjectively wrong", it will.

Beyond this clarificatory worry, there are replies that defenders of the wrong-making view** can provide to make the conclusion above appear less radical than

¹⁰⁸ For good critique of this view of excuses, see Bruno (2022). For an alternative view of excuses, see Strawson (§1.1).

¹⁰⁹ Similarly, the fact Andy lacked volitional control over his action would probably also excuse him "too much" – it would show that he is not blameworthy and that his action is not wrong (because "ought implies can").

it may initially seem (assuming that “wrong” in *standard view of excuses* refers to “subjectively/perspectively wrong”). Below, I briefly focus on one such reply.

The *standard view of excuses* is concerned with excuses that leave wrongness in place but entail that the agent is not at all blameworthy for the act or omission – i.e., *full excuses*. But there are also *mitigating* or *partial excuses* – i.e., factors that bear on what kind of blame response, or intensity of blame, is fitting *vis-à-vis* a blameworthy act or omission.¹¹⁰ Stress, lack of good sleep, being very much in love, might be such factors. To illustrate, suppose Elin deliberately lied to Matilda and that it was wrong for Elin to lie to Matilda. Matilda blames Elin for having lied to her and she acts overtly on one of the desires her blame has prompted – e.g., she acts overtly on the desire to demand an apology from Elin. Elin replies that she felt stressed when lying to her, and she believes that that reply will annul her blameworthiness. It does not. Elin’s reply does not show that she did not lie deliberately. Nor does it show that she was not at the moment of acting aware of the facts, such as the fact that they have a deal to be open and honest with one another, that constitute normative reasons which, taken together, make it wrong for her to lie to Matilda. However, her reply might well excuse her either in the sense that it makes further overt blame unfitting – it is not that important for Matilda to blame Elin overtly any more for a “minor” wrongdoing that she performed while stressed – or in the sense that it decreases the intensity of blame it is fitting to direct towards Elin. The wrong-making view** puts no pressure on such kinds of excuses.

Before moving on, it is appropriate to say a few words about moral ignorance and whether it excuses an agent. Consider Andy in §3.3 again. Suppose that, in response to being asked to explain his action, he claims that he did not believe that bullying Dwight was wrong. Further, let us suppose also that he speaks truthfully and that we cannot trace his moral ignorance back to some willing and witting wrongdoing. Is he then excused, on my view?

I believe that Andy’s reply does not excuse him in any way. However, I believe that knowledge (broadly construed) of the deontic status of one’s action or omission works as an *aggravating circumstance*. That is, it is a factor that increases

¹¹⁰ For more on how excuses can make overt blame unfitting and mitigating, or partial, excuses, see Woods (2021) and Sliwa (2020).

the intensity of blame it is fitting to direct upon the blameworthy agent, which is to say it modifies her degree of blameworthiness (recall §1.2.1).

Defenders of the witting wrongdoing view might wonder *why* it is fitting to blame Andy for bullying Dwight (given that he did not believe that it was wrong to bully him and that his moral ignorance could not be traced back to some blameworthy action or omission etc.).

Recall the assumption I made at the beginning of this chapter: I assume that the agents I will be talking about are (*morally*) *responsible agents*. That is, they have the capacity to assess (both “moral” and “non-moral”) normative reasons for and against certain actions or omissions, to modulate their behaviour in light of their awareness of these reasons, and so on. In addition, I assume that they are not in some form of extreme circumstances at the moment of acting or omitting that would hinder their abilities or capacities from working properly, such as suffering from sudden paralysis or being a victim of a major earthquake.

In addition, recall that in §3.5 I stated that “aware” in (ii) refers to facts the agent is either conscious of or facts that are self-accessible to her – i.e., she can know, recognise or tell the facts solely by means of reflection and self-examination. For example, I am aware of the fact that it is your birthday today either if I am conscious of it or I can easily come to recognise that it is your birthday today by means of reflection.

Now, my brief reply to defenders of the witting wrongdoing view is as follows: I believe that it is fitting to blame Andy for bullying Dwight because Andy *could and should* have believed that bullying him is wrong. Given his reasons-responsive capacities, given that he was not in any extreme circumstances, and given that he was aware of the relevant facts, such as the fact that his action will probably cause Dwight significant suffering, he could and should have believed that bullying Dwight is wrong.¹¹¹ His epistemic performance fell below an epistemic standard we reasonably could apply to him. Therefore, it is fitting to blame him.¹¹²

¹¹¹ This reply is capacitarian in virtue of the “could and should” claim. For more about capacitarianism, see Clarke (2017), Rudy-Hiller (2017), Ayars (2021), and Murray (2017).

¹¹² One might agree with me that Andy merits some form of blame. However, not blame as I understand it. According to Björnsson, failures such as Andy’s merit blame only when they are a result of the agent’s substandard quality of will. An agent’s quality of will, on his view, is “a matter of how well she cares about what is morally important” (Björnsson 2017: 149). Failures that are not explained by a fault in the agent’s cares, or quality of will, are not something for

Given this, we can improve the wrong-making view** like this:

*The wrong-making view****: an agent A is blameworthy for having performed a certain action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) A ϕ -ed despite, at the time of action or omission, being aware of the fact(s) that constitutes a normative reason(s) which, taken together, makes it wrong for A to ϕ , and (iii) A could and should have believed that ϕ -ing is wrong.

3.6.3 Degrees of Blameworthiness

The discussion in the precious section connects with the issue of how to understand what makes agents more or less blameworthy for having performed certain actions or omissions.

It is not evident how we can use the wrong-making view*** to explain degrees of blameworthiness.¹¹³ Consider an agent, A , who deliberately lied to someone, who was aware of the relevant facts when acting, and who could and should have believed that it is wrong for her to lie. And consider another agent, B , who deliberately killed someone, who was aware of the relevant facts when acting, and who could and should have believed that it is wrong for her to kill. Intuitively, we think A is less blameworthy than B . However, both A and B seemed to be equally aware of the relevant facts, have equally much control over their actions, and so on. So, appealing to the three necessary conditions included in the wrong-making view*** does not seem helpful when trying to settle the question of degrees of

which she fittingly can be blamed. Such failures can merit “skill blame”, where this is something different from, and less severe than, blame as I understand it – e.g., it does not involve being disposed to resent the agent, or feel guilt over one’s failure. In short, even if we stipulate that Andy is a decent person and does not harbour any flawed care for Dwight, and such like, I do not share Björnsson’s feeling that blame, as I understand it, seems unfitting to direct upon Andy, especially Dwight’s. For other blame scholars that seem to share my feeling, see Ayars (2021) and Clarke (forthcoming).

¹¹³ One might claim that defenders of the wrong-making view*** can appeal to degrees of wrongness – i.e., the “more” wrong my action is, the more blameworthy am I for having performed it. In short, I do not consider that way because it is controversial. Deontic concepts like “wrongness” are commonly thought to not admit of degrees, in contrast to evaluative concepts that are taken to admit of degrees.

blameworthiness. The aim of this section is to sketch a way for defenders of the wrong-making view*** to explain degrees of blameworthiness.

In §3.5, I presented the distinction between goodness and the desirable, on the one hand, and badness and the undesirable, on the other. For example, an action or omission can be the right one to perform, or omit, but bad. Now, according to one suggestion, we should consider the badness of actions or omissions when we are trying to explain differences in degree. For simplicity, let us grant that we measure the badness of an action or omission by appealing to the amount of harm it produced, and let us understand harm in terms of the amount of suffering produced. The more suffering it produced, the more blameworthy the agent is for having performed the relevant action, or for having omitted to perform the relevant action. The less suffering it produced, the less blameworthy the agent is for having performed the action, or for having omitted to perform the action.¹¹⁴

At face value, this suggestion might seem appealing, but I do not favour it. Suppose two agents – *C* and *D* – intentionally perform the same type of act, say, lying, and are equally aware of the facts that constitute normative reasons which, taken together, make it wrong for them to lie, and so forth. Further, their acts produce equal harm. However, *C* lies with the motive of causing unnecessary suffering and *D* with the motive of thereby impressing a person she is in love with. Intuitively, *C* is more blameworthy than *D*. But this cannot be explained by appeal to the harm the actions produce since they produce equal harm.

Not only that, but the suggestion allows for resultant moral luck. Resultant moral luck occurs when an agent performs an act or omission with a result that is (at least) partly beyond her control, and that result affects her moral status, in this context, her degree of blameworthiness. Consider *C* and *D* again. However, now suppose that they lied with the same motive – to cause unnecessary suffering – and that *C*'s target felt devastated in response but *D*'s target felt nothing in particular. According to the suggestion under discussion, *C* and *D* are blameworthy to different degrees: *C* is more blameworthy than *D*. This result does

¹¹⁴ This suggestion echoes our current practice of punishment a bit. As Nagel has observed, “[...] the penalty for attempted murder is less than that for successful murder – however similar the intentions and motives of the assailant may be in the two cases.” (Nagel 1979: 29) It should be noted, however, that we also punish people differently depending on the quality of their motives – e.g., the penalty for man-slaughter is less than for murder. Thank you, David Alm and Jakob Werkmäster, for reminding me about these practices.

not strike me as correct. I think that *C* and *D* are blameworthy to the same degree. It was mere luck that one of the targets felt devastated and the other did not. Such luck should not affect their degrees of blameworthiness.¹¹⁵

In the previous section, we learned that the agent's knowledge (broadly construed) of the deontic status of her act or omission counts as an aggravating circumstance – i.e., it is a factor that increases the intensity of blame that it is fitting to direct upon the blameworthy agent. We also saw that stress and lack of good sleep are considerations that count as mitigating circumstances – i.e., factors that decrease the intensity of blame it is fitting to direct upon the blameworthy agent. I believe that the lesson to draw from this is that bad motives count as aggravating circumstances – i.e., factors that increase the intensity of blame it is fitting to direct upon the blameworthy agent. In what follows, I will assume that all of these factors are relevant when we are assessing an agent's degree of blameworthiness.¹¹⁶

In sum, one can interpret wrong-making view*** as setting a *threshold* for blameworthy conduct: when an agent has satisfied the conditions stipulated by the wrong-making view***, she is blameworthy and a certain intensity of private blame is fittingly directed upon her. However, as we have learned in this and the previous section, there are factors that determine whether we should increase or decrease the intensity of blame that is fitting, i.e., that affect her degree of blameworthiness. For example, if she knew her action was wrong, more intense blame is fittingly directed upon her. Spelling out this theory in more detail – and, specifically, explaining more precisely how we should understand “threshold” – is a task better left for future research.¹¹⁷

¹¹⁵ For more on degrees of blameworthiness, resultant moral luck and why results should not affect agents' degree of blameworthiness, see Enoch and Marmor (2007) and Zimmerman (1987).

¹¹⁶ It might be that difficulty, or how hard it was for the agent to act rightly, are other necessary components. Due to space, I have not considered this option. For more about difficulty and degrees of blameworthiness see, Nelkin (2016).

¹¹⁷ This view is similar to, although also significantly different from, Miller's (forthcoming) view of blameworthiness. According to him, roughly, the witting wrongdoing view (or the wrong-making view as I have construed it) sets a *threshold* for blameworthy conduct, and an agent's quality of will affects her *degree* of blameworthiness.

3.7 Conclusion

The main question of this chapter was: what makes it fitting for anyone to blame an agent for having performed a particular action or omission? Apart from first stating how my discussion in this chapter relates to the distinction between attributability and accountability and the distinction between *de re* concern and *de dicto* concern, I presented two cases any theory about what makes agents blameworthy needs to account for. Then I analysed two accounts of what makes agents blameworthy for having performed particular actions or omissions in terms of how well they accounted for the cases. I argued that neither could adequately account for the two cases. Instead, I proposed a modification of one of the accounts – the wrong-making view** – and claimed that it could account for the two cases satisfactorily. Thereafter, I briefly defended and developed it further by considering some other challenges to it.

As is perhaps clear from this chapter, the discussion about what makes agents blameworthy is difficult. It connects to several large and difficult debates, such as the debate about what control is and what we can control, what it is to be aware of facts, how we should understand the deontic concepts “rightness” and “wrongness”, what capacities an agent must have to be a (morally) responsible agent, to name four debates. In this chapter, I have provided a *sketch* of what makes it fitting for anyone to privately blame an agent for having performed a certain action or omission. We can use this sketch to make the FA analysis of blameworthiness presented in Chapter 1 (§1.2) more informative and to make the views I will discuss in Chapter 5 more informative (there I will consider views that state that an instance of blaming a blameworthy agent *B* is proportionate if, and only if, its degree matches *B*’s degree of blameworthiness). However, for this sketch to turn into a good account of what makes agents blameworthy, more needs to be said its details (e.g., what is meant by “aware” and “could and should”) and the challenges facing it (e.g., the challenge from blameworthy right doing).

4 On Standing to Blame

In Chapter 1 (§1.2) I said that, in Chapter 6, I would test the FA analysis of blameworthiness – an analysis which states that an agent *A* is blameworthy for having performed particular act or omission ϕ to degree *n* if, and only if, it is fitting for anyone to blame *A* for having performed ϕ to degree *n* – by appealing to central claims made in the debate about the ethics of blaming. So, before I move on to test the FA analysis of blameworthiness, it will be helpful to know more about the norms concerning standing to blame and proportionality. In this chapter, I focus on standing to blame. In the next, I will focus on proportionality.

Although much has been written about the norm concerning standing to blame, there are still important unexplored issues. In this chapter, I will focus on one such issue.

Most scholars have focused on standing to blame someone *overtly*. At the same time, the majority view is that we also need to have a certain standing in order to blame someone *privately*. Consider, for example, the following selection of quotes:

But as we have seen, standing is broader than this. One can lack the standing to blame even privately or to a third-party. (Fritz & Miller 2022: 771-2)

Our relationship to the agent is relevant to determining not only whether we have standing to express moral criticism to her for an objectionable action or attitude, but also whether we have standing to adopt particular sorts of blaming attitudes towards her. (Smith 2007: 479)

The thought that conditions of standing also apply to private blame is intuitive. Suppose that an unrepentant cheater, *A*, in response to hearing of another agent, *B*, cheating on *C*, starts to privately blame *B* for cheating on *C*. Even if *A* does not blame *B* overtly, there is something morally problematic about *A*'s private blaming of *B*.

One explanation of this is that *A* lacks the standing to privately blame *B*, the idea being that doing something one lacks standing to do is *pro tanto* wrong. In this chapter, I argue that this explanation should be resisted. I argue instead that many key features of standing to blame do not apply to private blame. For example, private blame is not the exercise of some normative power and *A*'s private hypocritical blaming of *B* is not *pro tanto* wrong. Thus, contrary to the prevailing view, it is doubtful that there are conditions of standing to private blame.

Instead of appealing to the conditions of standing to explain what is morally problematic about *A*'s private blaming of *B*, I follow Matt King (2019a; 2019b) and argue that what is morally problematic here is that *A* has violated a norm concerning priority and attention: *A* ought to attend to her own wrongdoing before attending to that of others.¹¹⁸

Importantly, my conclusions in this chapter do not generalise to overt blame. In other words, my conclusions do not justify full-blown scepticism about the standing to blame.¹¹⁹

The chapter is structured as follows. First, in §4.1, I present the distinction between private and overt blame that I presented in Chapter 2 (§2.8-§2.9) and highlight a key difference between them. Then, in §4.2, I offer an overview of the idea of standing to blame and reveal some key features of it, including the idea that standing is a right of some sort. In §4.3, I argue that most of the key features do not apply to private blame in a way that would justify us in thinking that there are norms of standing to blame someone privately. I end with some concluding remarks in §4.4.

¹¹⁸ Granting that the stakes are not so that *A* has a conclusive reason to attend to *B*'s wrongdoing. From now on, when I refer to King's view, I will assume that agents such as *A* has a conclusive reason to attend to their own wrongdoing before attending to that of others.

¹¹⁹ King (2019a) argues for such a broad sceptical claim.

4.1 Private Blame and Overt Blame

Recall from Chapter 2 (§2.8-§2.9) that by “private blame” I mean that one harbours the sentiment of blame. Importantly, harbouring the sentiment of blame does *not* involve interacting with its target in any way. Nor does it involve the performance of some overt act. The blame sentiment can manifest itself in a desire to demand an apology from a particular person. However, one does not need to act on that desire. In contrast, when we (sincerely) blame someone overtly, we have the blame sentiment and act on its manifestations. For example, we act on the desire to demand an apology.

One might argue that the sentiment of blame – when manifested – always involves some involuntary observable physical manifestations such as a higher pulse rate, blushing, and so on. Hence, it would be an exaggeration to say private blame involves no interaction with its target and no overt actions of any kind, since private blame is never completely “private”, especially when it is manifested.

I accept that the sentiment of blame, when it is manifested, may involve involuntary, observable physical manifestations, like blushing. However, as will be recalled from §2.6, I believe that blame does not always involve such manifestations. In addition, I believe that blame, when it has such clear manifestations, that we can be good at hiding those manifestations and that in many cases it is hard to tell whether or not someone has a given sentiment. For example, it may be very difficult to see whether someone sweating in the presence of another person indicates that she blames her, hates her, envies her, or something else – or indicates nothing at all for that matter: maybe she is just hot.

Further, what I have in mind by overt blame is something voluntary, not something involuntary, such as blushing or sweating.

Having briefly clarified my usage of private blame and overt blame, I turn to stress a further way in which they differ. This difference will be relevant later when I argue for scepticism about the standing to blame someone privately.

As will be recalled from Chapter 3 (§3.4), actions, such as overt blaming, are things we generally have *volitional control* over. By that I mean, roughly, that we can blame someone overtly willingly, or for more or less whatever reason we think is sufficient – e.g., to win a bet or to impress someone. In addition, we can choose when we want to blame someone overtly – now or later. The overt blame is in

this sense “up to us”: I can see to it whether it is or is not performed. In contrast, attitudes like private blame are things we generally do not have volitional control over. In the private case, we cannot blame someone deliberately, or post-pone blaming her. Nor can we blame, or cease to blame, *for* more or less any pragmatic or prudential reason, including in order to win a bet or impress someone. Rather, just as our beliefs are only responsive to (if to anything) what we take to be epistemic reasons, our private blame is only responsive to what we take to be “the right kind of”, or “standard”, reasons for private blame.¹²⁰

The mere fact that we do not have volitional control over our attitudes, and *a fortiori* over private blame, need not be taken to imply that these attitudes completely escape our control. As stated in Chapter 3 (§3.4), we presumably have rational control over our attitudes. By “rational control” I meant the sort of control we exercise by being receptive and reactive to certain reasons – that is, by forming, revising, etc., our attitudes in light of our awareness of certain facts. In connection with this, it makes sense to ask someone to justify, or explain, *why* she is afraid of something, *why* she believes something, or *why* she blames someone privately. The application of such “why-questions” indicates that we hold these kinds of attitude and emotion for some reason.¹²¹

Finally, note that we can trigger our sentiment of blame willingly. For example, if I have acquired the blame sentiment about a particular agent, I can trigger it by, for instance, deciding to meet the blameworthy agent to discuss her wrong action. Relatedly, we can decide to change our circumstances so that, as a result, we acquire new attitudes. For example, I can decide to investigate my own past behaviour. As a result, I might start to blame myself privately because I realise that I have on some occasion deliberately acted despite being aware of the reasons that, together, make it wrong for me to perform that action or omission. This is similar to the way in which we can acquire the belief that the light is on by deliberately switching the light switch.

¹²⁰ For more on this claim, see Heuer (2010) and Hieronymi (2005).

¹²¹ For more on why the application of “why-questions” is a decisive criterion of reason-responsiveness, see Chapter 2 of Marušić (2022).

4.2 Core Features of Standing

Here I present some key features of the idea of standing to blame. I will not specify what “blame” refers to. However, in the next section, I will argue that most of the features I have detailed do not apply to private blame in a way that justifies us in thinking that there are norms of standing to blame someone privately. Before I do that, I introduce the general idea of what it is to have standing to blame.

It is usual to introduce the idea of standing to blame through some version of the following imagined conversation:

A: I blame you for φ -ing.

B: Who are you to blame me for φ -ing?

When agent *B* utters “Who are you to blame me for φ -ing?” she is not questioning agent *A*’s verdict concerning her blameworthiness for φ -ing. *A*’s verdict might very well be correct. *B* may even believe this. Rather, what *B* is questioning is *A*’s standing to blame her, or *A* being in a proper position to blame her. Intuitively, there is such a thing as being more (or less) well-placed or ill-placed to blame someone. Gerald A. Cohen (2006) says something similar here:

We can distinguish three ways in which a person may seek to silence, or to blunt the edge of, a critic’s condemnation. First, she may seek to show that she did not, in fact, perform the action under criticism. Second, and without denying that she performed the action, she may claim that the action does not warrant moral condemnation, because there was an adequate justification for it, or at least a legitimate excuse for performing it. Third, while not denying that the action was performed, and that it is to be condemned (which is not to say: while agreeing that it is to be condemned), she can seek to discredit her critic’s assertion of her standing as a good faith condemner of the relevant action. (Cohen 2006: 119)

Moving to the core features, most theorists of standing agree that the nature of standing is a *right* of some sort. That much is uncontroversial. What is controversial is how we should specify the right. On one suggestion, the standing should be identified with having a *privilege-right* to blame someone, and an agent *A* has a privilege-right to φ if, and only if, *A* has no obligation not to φ .

Another suggestion is that standing should be identified with a *power-right*, i.e., with having the proper authority to exercise a *normative power*. Normative powers change the agent's normative landscape – e.g., by creating, undercutting, intensifying, or attenuating her reasons. One needs to have a certain authority to successfully exercise such normative powers. If agent *A* does not have the standing to blame agent *B*, *A* cannot give *B* a *pro tanto* reason to (say) explain her behaviour. That is the rough idea.

One argument against understanding the nature of standing as a power-right is presented by King (2019a). It is helpful to consider this argument here so that we can better understand what we might call the power account.

According to King, *A* cannot alter *B*'s normative situation *at all* by blaming her. *B* has a *pro tanto* reason to apologise for her action, or some such thing, in virtue of being blameworthy for having performed it. However, as Kyle G. Fritz and Daniel J. Miller (2022) point out, King's dismissal seems too quick. They argue that although before being blamed by *A*, *B* might have a *pro tanto* reason to apologise for her action, or some such thing, *A* can provide *B* with an *additional pro tanto* reason to apologise for her action. I believe it is more plausible to think that *A* can *intensify* *B*'s *pro tanto* reason to apologise for her action, or make it more *urgent* that *B* apologises for her action *now* rather than later, instead of thinking that *A* gives *B* a further *pro tanto* reason to apologise.¹²²

It is common to adopt mixed views of what standing is – in the sense that the analysis combines privilege-rights and power-rights. For example, James Edwards argues that:

A has standing to hold *B* responsible for *X* by doing *Y* if and only if *A* has both: (i) An agent-relative privilege to hold *B* responsible; (ii) The power to put *B* under a duty – by holding *B* responsible – to offer a fitting content-sensitive reply to *A*. (Edwards 2019: 456)

Fritz and Miller (2022) propose another kind of mixed view. I discuss this in the next section.

Another central idea is that *specific conditions* determine whether or not an agent has the standing to blame someone. Several of these conditions are mentioned in

¹²² Thank you, Benjamin Kiesewetter, for mentioning some of these options to me.

the literature. Here, I focus on the two most common. In the next section, I will argue that none of the conditions clearly applies to private blame.

- The avoidance of meddling (or “none-of-your-business”) condition.
- The absence hypocrisy (or “non-hypocrisy”) condition.

In cases of meddlesome blame, *B* points to facts showing that *A* is an outsider with regard to the underlying matter at hand in which she is trying to intervene. Being an outsider as regards the matter at hand is what undermines *A*’s standing to blame *B*. Familiar responses to blame performed under such circumstances are “This has nothing to do with you!” and “Mind your own business!”

However, the most frequently discussed condition is the second one: absence of hypocrisy condition. There is a lively debate about what hypocritical blame, and more generally hypocrisy, is, and why it would undermine standing. I will limit the discussion to the two leading theories – *the commitment account* and *the moral equality account*.¹²³

On both views, roughly, hypocritical blame is essentially about unjustified exception-making with respect to a certain (moral) norm *N*. The hypocritical blamer engages in (or, perhaps, is prone to) blaming others for violations of *N* that she herself committed without also subjecting herself to blame, and without a justifiable reason for this.¹²⁴ For example, according to Fritz and Miller (2018),

[...] being hypocritical with respect to blame is a matter of having an unfair differential blaming disposition (UDBD) – having a disposition to blame others for a violation of some norm *N* but lacking a disposition to blame oneself for violations of *N* without having a justifiable reason for this difference. (Fritz & Miller 2018: 122)

In many cases, such blame indicates a lack of genuine commitment to *N* on the blamer’s part. It is this lack of commitment to *N* that undermines her standing to blame others for violations of *N*, argue defenders of the commitment account (Crisp & Cowton, 1994; Todd, 2017). In contrast, defenders of the moral equality account argue that by making exceptions for herself (or others), the

¹²³ Both proposals have their pros and cons. For criticisms of the commitment view, see Fritz and Miller (2019). For criticisms of the moral equality view, see Lippert-Rasmussen (2020).

¹²⁴ Importantly, the violations need not be identical; they need only be sufficiently similar.

blamer (implicitly) rejects the equality of persons – i.e., the notion that morality applies equally to us all – with respect to *N*. It is in virtue of this that her standing to blame others for violations of *N* is undermined (Fritz & Miller, 2018; Wallace, 2010).¹²⁵

Familiar responses to hypocritical blame are “You of all people can’t say that!” and “Look who’s talking.”.

Another core feature I wish to mention is that standing-less blame is *pro tanto* wrong, not only *pro tanto* wrong in general, but *pro tanto* wrong as far as the blameworthy agent is concerned as well.¹²⁶ As Ori Herstein (2020) puts the same point:

[...] to intervene without standing is to wrong the party whose valuable interests gave rise to the duty not to intervene in the first place. Accordingly, breaching standing’s duties of nonintervention is not only wrongful simpliciter, it is also a wrong to those intervened with. (Herstein 2020: 12)

We can support this claim by attending to our practices of blaming and criticising. Some kind of moral criticism is often considered a fitting response to, for instance, meddlesome or hypocritical blaming. Further, if we were to meddle in someone’s affairs, we would probably apologise for the intervention or otherwise explicitly acknowledge that we are in the wrong – “I know it is wrong of me to intervene in your affairs, but...” This is indicative of our intuition that standing-less blame is *pro tanto* wrong.

The final key feature I wish to mention is that conditions of standing should *guide* us. That is, they should help us to blame others responsibly. (What use is there for conditions of standing if they cannot be followed?) For example, if one learns that one’s blaming of another is rightly regarded as meddlesome, then one should be able to respond to that fact by ceasing to blame.

¹²⁵ For a relatively similar view about hypocrisy and standing, see Roadevin (2018).

¹²⁶ Cohen does not seem to agree with this claim, however. Consider what he writes in the following footnote: “My topic is not when it’s morally permissible or obligatory to condemn, and it is not part of my view that it is always bad or wrong for someone who is not in a position to condemn to condemn” (Cohen 2006: 119, fn. 10). It might be that Cohen agrees with the claim that standing-less blame is always *pro tanto* wrongful. What he objects to is that it is always *all-in* wrong.

4.3 Scepticism About the Idea of Standing to Blame Someone Privately

In this section, I aim to show that (i) some explications of “standing” entails that there is no such thing as having (or lacking) standing to blame someone privately (§4.3.1); (ii) that the avoidance of meddling condition is not applicable to private blame (§4.3.2); (iii) that private hypocritical blame is not *pro tanto* wrong and that it is therefore questionable that there is a standing condition regarding hypocrisy for private blame (§4.3.3); and (iv) that the conditions of standing do not guide private blamers adequately (§4.3.4).

4.3.1 The Nature of Standing and Private Blame

In §4.2, I presented some common ways of understanding the nature of standing: as a privilege-right, a power-right, or a mix of both. If one thinks that standing necessarily involves a power-right in all cases, standing does not apply to private blame. We do not exercise a normative power by blaming someone privately. For example, I do not give someone a *pro tanto* reason, or intensify someone’s *pro tanto* reason, to apologise for her action merely by having the sentiment of blame.

Fritz and Miller (2022) take the observation that we do not exercise a normative power by blaming someone privately to be a good reason for concluding that the nature of standing cannot solely be, or necessarily be in all cases, identified with a power-right. Consider what they write here:

Nevertheless, the standing to blame cannot solely be a power, as Edwards acknowledges [see §4.2]. Edwards sees standing involving both a privilege and a power. There is good reason for this. Edward is focused on directed blame, or blame that is expressed and directed toward the individual blamed. But as we have seen, standing is broader than this. One can lack the standing to blame even privately or to a third-party. In such cases, the offender might be unaware of the blame. If so, it’s implausible that the blame imposes obligations on her. Thus, the standing to blame can be understood as (or as required for) a privilege-right in all cases of blame, whether private, overt, directed, or non-directed. But it might also be understood as (or as required for) a power-right in cases of directed, overt blame where the individual wronged blames the wrongdoer, thereby creating certain obligations for her. (Fritz & Miller 2022: 771-2)

Their reason for thinking standing should not be identified solely, or necessarily in all cases, with a power-right is that standing is “broader than this”. According to them, standing also applies to private blame and third-party blame.

Since in this chapter I am focusing on private blame, I do not take a stance on the feasibility of the claim that there is such a thing as having, or lacking, the standing to blame a third-party. However, showing that there seems to be no such thing as standing to blame someone privately, I weaken Fritz and Miller’s argument by undermining the plausibility of one element of it. Ultimately, they might very well be correct in that standing cannot solely refer to a power-right given the phenomenon of third-party blame. That, however, is a question for further investigation that I will not undertake here.

4.3.2 Conditions of Standing and Private Blame

Suppose you witness a stranger confessing to their partner that they have had an affair. The partner starts to blame the cheater overtly. You have no relationship with either party. Nor are you injured or affected in other ways by the stranger’s action. In this case, I have the intuition that you lack standing to blame the stranger overtly but do not lack standing to blame her privately. What this case suggests is that the avoidance of meddling condition is only a condition of overt blame.¹²⁷ Patrick Todd (2017) seems to share this feeling when he writes:

As I see it, someone’s wrongdoing being “some of one’s business” is a condition on expressing blame, whereas the fact that one’s blame would be “hypocritical” is a condition on even (so to speak) feeling blame. (Todd 2017: 348)

Fritz and Miller, and others, might agree with me that the avoidance of meddling condition only applies to overt blame. This, however, does not prohibit them from claiming, as Todd – a proponent of the commitment account (§4.2) – does that hypocrisy is a norm of standing that can undermine both standing to blame someone overtly *and* standing to blame someone privately.

¹²⁷ Cohen’s claim that “[...] much of what disqualifies the act would also disqualify the attitude, and that, as it also seems to me, a major reason why the act gets disqualified, in the relevant cases, is that it expresses a disqualified attitude” (Cohen 2006: fn 31) must therefore be false. For in cases of meddlesome blame, it is not true that the attitude is disqualified.

Next, I consider some reasons challenging the claim that hypocrisy undermines one's standing to blame someone privately. Given that hypocrisy is the standing condition *par excellence*, my thought is that if absence of hypocrisy is not a condition of standing to blame someone privately, then we have good reason to be sceptical about the idea that conditions of standing apply to private blame altogether. In other words, if the two most obvious conditions pertaining to standing – the avoidance of meddling condition and the absence of hypocrisy condition – do not apply to private blame, it is very unlikely that there is such a thing as standing to blame someone privately.¹²⁸

4.3.3 Hypocritical Private Blame and Wrongness

Suppose agent *A* witnesses another agent *B*'s act of infidelity and, in response, directly starts to blame *B* privately for that reason. *A* has committed a similar act herself without making proper amends. *A* does not blame herself privately for the infidelity, and she lacks a justifiable reason for this. Even if someone were to tell *A* that she lacks the standing to blame *B* privately, she could not cease from blaming *B* directly for that reason. The only reason that would work would be a reason showing that *B* has made proper amends or did not, in fact, cheat. But that is not the case here. Thus, it is hard to see that *A* does anything *pro tanto* wrong by blaming *B* privately in this case, as *A* cannot help but blame *B* privately given her knowledge of *B*'s infidelity.

There might be a worry that my argument above (that *A*'s private hypocritical blame of *B* is not *pro tanto* wrong because *A* lacks volitional control over her private blame) is merely an application of a more general scepticism about doxastic and attitudinal wrongdoing – that is, about the idea that beliefs and attitudes can be wrong or wrong others.¹²⁹

While general scepticism about doxastic and attitudinal wrongdoing would be sufficient for my argument, I need not argue for such a broadly sceptical view here. For my purposes, a more modest claim suffices: an agent neither does what

¹²⁸ Or the conditions for having standing to blame someone privately are different from the conditions for having standing to blame someone overtly.

¹²⁹ There is a lively debate as to whether attitudes and beliefs can be wrong or wrong others. Compare: Sher (2021), Basu and Schroeder (2019), and Enoch and Spectre (forthcoming).

is wrong nor wrongs anyone when she has beliefs or attitudes that are adequately justified or fitting.¹³⁰

As I have set up the case about *A* and *B* above, *A* is adequately justified in adopting her private blame of *B*. *B* is blameworthy and it is fitting to blame *B*. *A* privately blames *B* in response to learning, let us grant sufficiently, about *B*'s infidelity. There is no evidence, no reason, that we can provide *A* with that shows her attitude towards *B* to be unfitting – e.g., we cannot point out to *A* that *B* did not cheat on her partner *C* – nor can we argue that *A* blames *B* without having gathered sufficient evidence of the circumstances because she has. Therefore, it cannot be argued, without violating the intuitively modest claim stated above, that *A* does wrong, or wrongs, *B* in blaming her privately.

That said, *A*'s attitude towards herself does exhibit some form of reasons-unresponsiveness: she is not blaming herself privately for her cheating even though she, presumably, at some level, knows that she has cheated. Some form of criticism is fittingly addressed at *A* as regards her attitudes towards herself. And we could remind *A* about her infidelity or give her more evidence about her infidelity which, if *A* were reasons-responsive, etc., would result in her blaming herself privately.

In reply to these worries from (what we can call) control and rationality, one might argue that the fact that we sometimes feel *guilt*, *regret*, or some sort of *dissonance* when we blame someone privately and hypocritically point to the fact that private hypocritical blame is wrong.¹³¹

We often regret things that are not wrong. And we often feel guilt, or some sort of dissonance, over things that are not wrong. Consider, for example, the following version of Williams' famous lorry driver case: a lorry driver feels guilt over having run over and killed a child that suddenly jumped in front of his truck while he drove. The driver's guilt does not show that he did anything wrong by hitting the child with his truck (as he could not avoid that). Likewise, the fact that I feel guilt over being the only survivor in a severe earthquake need not imply that I have done something wrong by surviving. To the point: the presence of feelings

¹³⁰ According to Schroeder and Basu (2019), who are leading proponents of doxastic wrongdoing, in all instances of doxastic wrongdoing the agent holds epistemically unjustified beliefs.

¹³¹ Thank you, Justin Snedegar and András Szígeti for presenting this argument to me.

such as guilt and dissonance do not always show we have done something wrong. Therefore, the above reply is not convincing.

Hypocrisy is said to undermine one's standing, one's privilege, to blame someone. Doing something one lacks the privilege to do is to do something *pro tanto* wrong. However, as I have shown in this section, blaming others privately and hypocritically is not *pro tanto* wrong. Therefore, either absence of hypocrisy is not a condition of standing for private blame or what I have said above is incorrect and *A* is in fact doing something *pro tanto* wrong by privately blaming *B*. Given the arguments presented in this section, I think the first option is the more plausible one. At the very least, proponents of the claim that *A* does something she lacks standing to do in blaming *B* need to say something more.

Until now, I have been content to present some reasons why we should be sceptical about the claim that there is such a thing as standing to blame someone privately, with a focus on private hypocritical blame. I have not yet said anything about the case presented at the beginning of the chapter that drove the intuition that there is such a thing as standing to blame someone privately. Below, I offer an alternative explanation of what is morally problematic in that case – one that does not invoke the concept of standing to blame.

My claim that merely harbouring private hypocritical blame is not *pro tanto* wrong is not inconsistent with asserting at the same time that there is something morally objectionable about private hypocritical blamers nonetheless. On my preferred explanation of the latter claim, private hypocritical blamers fail to prioritise the matters they are attending to adequately. In attending to *B*'s infidelity, *A* is attending to the wrong thing. She should focus on her own wrongdoing instead of others' similar wrongdoings – or, at least, *also* focus on her own wrongdoing.¹³² Presumably, we have a duty to prioritise attendance to our own past behaviour before attending to the behaviour of others. As King (2019a; 2019b) puts the same point, hypocritical blamers “[...] are failing to address their own faults, which ought to take priority over critiquing others. What the pot should do is polish itself” (King 2019a: 272). He continues:

¹³² If, by attending to her own actions, she then does not feel prompted to start blaming herself privately, it is questionable whether she has the relevant capacities to be a morally responsible agent on the whole (see the introduction to Chapter 3).

In the case of hypocritical blame, the blamer is critiquing others when their attention and efforts ought to be directed at improving their own conduct. Indeed, I argue that hypocritical blame is best understood as running afoul norms counselling improvement of one's own moral house, which I call norms of priority. (King 2019a: 273)

It will be recalled from §4.1 that we do not have volitional control over our attitudes. Rather, we have rather rational control over them. However, we do seem to have (at least some) volitional control over our attention. We can decide not to attend to “the mote in thy sister’s eye”. King provides a helpful example to illustrate the point. *A* is present at *B*’s wedding. During the festivities, *A* attends to *B*’s previous, yet minor, wrongdoings, triggering her, or acquiring, blame of *B*. In such a scenario, we might very well wish to say that it is not merely that the wedding is not the time and place to overtly address these pasts wrong; it is the wrong time and place for *A* to attend to *B*’s wrongdoing, and thereby to trigger her private blaming of *B* or start blaming *B* privately.¹³³

4.3.4 Guidance and Private Blame

As will be recalled from §4.2, the conditions of standing are supposed to guide us. That is to say, they should help us to blame others responsibly and judiciously. However, if they are to guide us, it makes sense to think that we can adjust our behaviour or attitudes in light of these conditions. For instance, it should be possible for *A* to refrain from blaming *B*, both overtly and privately, in light of, or because doing so would be or is (say) hypocritical or meddlesome. When it comes to private blame, as will be recalled from §4.3.3, *A* cannot cease to blame *B* privately for the reason that it is hypocritical, any less than she can cease to believe that it is raining outside for the reason that she will get a reward for doing so. Thus, it seems that the conditions of standing fail to guide us in cases of private blame.

In reply, one might argue that this conclusion is too quick. The observation suggests only that the specific conditions of standing cannot *directly* guide private blamers. One cannot directly cease to blame someone privately for, say, the reason it is hypocritical. But the same is true of many norms concerning beliefs and

¹³³ This is not to say that I agree with all of the claims King makes. Afterall, he is also a sceptic concerning standing. For example, I am unpersuaded by his argument that overt blame is not a normative power.

attitudes, such as the norm to avoid irrationality. What matters is that the norms can guide us *indirectly*.

If one has an irrational belief or unfitting attitude, one cannot directly change one's belief or attitude voluntarily for the reason that it is unfitting or irrational. What one can do is to reason, and expose oneself to evidence, in order to cease holding the irrational belief. The difference, of course, is that in the case of *A* and *B*, there are no reasons *A* could expose herself to such that *A* would cease to blame *B* privately. *B* is blameworthy and it is fitting to privately blame *B*.

I believe this difference is important. There is something unreasonable in the idea of norms of standing to blame guiding us to take indirect and atypical steps, such as taking a pill, undergoing hypnosis, or going to therapy, that desensitise us to reasons to privately blame in the name of being a responsible blamer. It is excessively demanding or, perhaps, unreasonably demanding.¹³⁴

The claim is not that morality cannot demand a lot from us. In the case of someone with racist attitudes, it is not excessively or unreasonably demanding to urge that she undergoes therapy in order to change her irrational attitudes, as, in this case, reasons are available that she could come to appreciate and in virtue of which she could change her attitudes. Similarly, it is not excessively or unreasonably demanding to insist that *A* should take steps enabling her to recognise that she, too, has cheated on her partner. My claim is merely that if norms of standing applied to private blame, those norms would sometimes guide us to take unreasonably demanding or intrusive indirect steps for the purpose of making us less reason-responsive.

In reply, one might perhaps argue that *B* is not blameworthy *relative to A* in virtue of *A*'s hypocrisy. And given that *B* is not blameworthy relative to *A*, there must be reasons of the right kind that *A* could respond to directly, and that would result in *A* no longer blaming *B* privately. In addition, it would not be unreasonably demanding to urge *A* to appreciate those reasons, and thereby cease blaming *B*.

¹³⁴ It is not particularly demanding to demand that *A* takes a walk around the house (to cool down her anger of *B*) or distracts herself. In short, as will be recalled from §4.1 and Chapter 2, by "private blame" I mean to merely harbour the sentiment of blame. Just because one succeeded in suppressing a manifestation of blame (e.g., an episode of resentment) is not the same as having ceased to blame.

privately. If *A* were to blame *B* nonetheless, *A* would be doing something unfitting – blaming someone who is not blameworthy relative to her.

In this thesis, I have been concerned with blameworthiness *simpliciter*, not *relative* blameworthiness (§1.2). It is therefore beyond the scope of this work to take a stance on the issue just mentioned about *relative* blameworthiness. Let me say something briefly about the above reply nevertheless.

Apart from the fact that we need to know how to formulate a plausible account of relative values (there is still no consensus regarding how we should understand relative values or agent-relative reasons¹³⁵), there is also a question about whether there is just *relative* blameworthiness or also blameworthiness *simpliciter*. If there just is *relative* blameworthiness and no blameworthiness *simpliciter*, we need to explain what is so special about blameworthiness, as similar thick value properties, such as “admirable” and “choiceworthy” are not commonly viewed as just relative values. If there is both kinds of blameworthiness, on the other hand, there is the question about what we should say about cases where *B* is not blameworthy *relative* to *A* but blameworthy *simpliciter*. In such scenarios, it seems fitting for *A* to blame *B* and not fitting for *A* to blame *B*, and one might wonder whether *A* in virtue of her hypocrisy lacks standing to privately blame *B* “*simpliciter*”.

Apart from these clarificatory concerns, if we accept the reply above, we need to reformulate the initial idea of standing to blame. As noted in §4.2, when *B* utters “Who are you to blame me for φ -ing?” she is not questioning *A*’s verdict concerning her blameworthiness for φ -ing. *A*’s verdict might very well be correct. *B* may even believe this. Rather, what *B* is questioning is *A* being in a proper position to blame her. If we accept the reply above, however, we need to say that *A* is not in a proper place to blame *B* and that *A* is making a factual mistake – it is not true that *B* is blameworthy (at least, not relative to *A*).

Moving to another worry about my initial argument, some might complain that my conclusion above is too quick for a further reason. Guidance, it might be said, is not only about ceasing to blame someone in response to the conditions of standing. It can also be about (say) apologising.¹³⁶ For example, one might argue

¹³⁵ For good critique of most suggestions of how to understand relative values or agent-relative reasons, see Schroeder (2007) and Bykvist (2018).

¹³⁶ Thank you, András Szígeti, for mentioning this aspect of guidance to me.

that the conditions of standing can guide private blamers in the sense of encouraging them to apologise for their hypocritical private blame, or something to that effect.

In the case with which I have been concerned, one agent privately blames another hypocritically, and someone appeals to the conditions of standing to discourage the first agent from such private blaming. I have focused on this kind of case because we have the intuition that it should be possible for *A* to refrain from blaming *B*, both overtly and privately, if reasons appealing to standing were to be relevant. Given that *A* cannot do this, or cannot do so in any plausible way, doubt arises over the very idea of standing to blame someone privately.

4.3.5 Summary

What, then, does the idea of standing to blame someone privately commit us to? Such standing cannot be a matter of having, or not having, a power-right. Blaming someone only privately is not exercising a normative power. So, if standing to blame is to apply to private blame it needs to be understood as a privilege-right. But does violating any of the specific conditions of standing entail a loss of the privilege to privately blame someone? I am doubtful that it does. The avoidance of meddling condition does not apply to private blame at all. Some have argued that the absence of hypocrisy condition is clearly applicable to private blame. However, as I have shown, merely blaming someone privately and hypocritically does not constitute a *pro tanto* wrong. However, the question whether one has a privilege right to privately blame or not would only arise if doing so could amount to a *pro tanto* wrong. So, we found that not even the absence of hypocrisy condition – arguably, the paradigmatic norm of standing to blame – applied to private blame. Finally, the conditions of standing were shown as not capable of guiding private blamers – at least, not in any plausible way that is not excessively demanding. I therefore strongly suspect that there is no such a thing as standing to blame someone privately.

Importantly, these findings do not generalise to overt blame. For example: we can exercise normative powers by blaming others overtly; the avoidance of meddling condition and absence of hypocrisy conditions seem to be clearly applicable to overt blame; and the specific conditions of standing can guide overt blamers

directly and in reasonable ways. It is, for instance, easy to refrain from blaming someone overtly for the reason that doing so would be meddlesome.

At the same time, I have also shown that one does not need to invoke the concept of standing in order to explain what is morally problematic about private hypocritical blamers. I can both deny the cogency of standing to blame privately and at the same time maintain that there remains something morally problematic about the agents in question. Specifically, on my preferred explanation, they have violated the norms of priority governing what they ought to attend to.

In sum, I have argued for scepticism about the standing to blame someone privately. Proponents of the claim that there is such a thing owe us a more concrete account, especially since there seem to be no declared norms of standing for other private attitudes, such as belief, fear, love and admiration. Until they have done that and answered my concerns, I am entitled to proceed on the basis that there is no such thing as standing to blame someone privately.

4.4 Conclusion

The question for this chapter was: Do norms of standing apply to private blame? I have argued for a sceptical position concerning the idea of standing to blame someone privately. Many key features of standing to blame do not apply to private blame. They would do so if there were such a thing as standing to blame someone privately.

There is more to be said about the details of my arguments. In particular, more should be said about what is morally problematic about hypocritical private blamers, but that is a task for the future.

5 On Proportional Blame

It is clear why common sense acknowledges a norm of proportionality. We tend to think that overdoing blame is unfair to the blameworthy agent and we also tend to think that underdoing blame is morally problematic too, among others because of the potential implications of this for the blamee's victim. As Alexander Edlich argues, blaming someone "too lightly, can be a way of harming a victim by denying them moral support in the face of wrongdoing" (Edlich 2022: 221). All the same, what exactly amounts to proportional blame is currently under-theorised.

Specifically, previous discussions have focused on cases involving a single blamer and a single blameworthy agent. As I stress in this chapter, more needs to be said about cases involving a plurality of blamers.

Further, insufficient attention has been given to the distinction between private and overt blame. I argue that, given their different character, and especially differences in the way they aggregate, different aspects will be relevant when assessing the proportionality of private and overt blame, respectively.

Finally, previous discussions seem to argue for a proportionality principle that echoes the condition applied in the domain of legal punishment or retributive justice – i.e., that the degree of legal punishment should match the degree of the crime or "fit the transgression". For instance, Neal A. Tognazzini and D. Justin Coates (2021) write in their entry on blame in *Stanford Encyclopedia of Philosophy*: "Analogous to the common thought that the punishment must fit the crime, it is plausible to suppose that the blame must, in some sense, fit the transgression". Consider also Fricker (2016):

[...] blame must of course be proportionate to the wrongdoing, for it is the degree of wrongdoing that justifies the degree of blame. It is not appropriate to be maximally censorious in respect of a small misdemeanour, though the typical scenario of blame (where the blamer is hurt) means it may often be tempting to allow oneself to do just that. (Fricker 2016: 168-9)

I argue that while such a “retributivist” view aligns well with private blame, it does not align well with overt blame. In short, this chapter aims to provide the beginnings of an adequate account of proportional blame.

Apart from it being of theoretical relevance to explore what it is for blame to be proportional, there is a practical imperative to understand the issue and to develop guidance on how to blame responsibly in scenarios involving more than one blamer. Given technological advances and how interconnected we are in today’s society, there has been a natural rise in the number of real-life cases involving a plurality of blamers in which people have arguably been blamed too much.¹³⁷

I will proceed as follows. First, in §5.1, I highlight some further differences between private and overt blame in addition to the differences already provided in Chapter 4. These insights will become relevant when later, in §5.2, I develop an account of proportional blame. In §5.3, I draw out some practical implications of my account. I end with some concluding remarks in §5.4.

5.1 Private Blame and Overt Blame Again

In Chapter 2 (§2.8-§2.9) and Chapter 4 (§4.1), I distinguished between private and overt blame. I claimed that to blame someone privately is to merely harbour the sentiment of blame and to (sincerely) blame someone overtly is to act on that sentiment’s manifestations – e.g., one may act on the desire to demand an explanation from the blameworthy agent by demanding one from her. I also highlighted another distinguishing feature of private versus overt blame: they involve different kinds of control – volitional control and rational control, respectively. In this section, I highlight further ways in which private and overt blame differ.

As the reader might recall from Chapter 1 (§1.2), we account for the ways in which affective attitudes and actions come in degrees differently. The standard approach to accounting for degrees of affective attitudes, such as sentiments,

¹³⁷ For good real-life examples of cases where people have overtly blamed someone too much, see (Billingham & Parr, 2020a, 2020b). Although Billingham and Parr speak of online public shaming, it is possible to interpret their cases as ones in which a plurality of agents have overtly blamed someone too much.

appeals to their intensity. That is, we admire Leonardo da Vinci's *Mona Lisa* more than Johannes Vermeer's *Girl with a Pearl Earring* when we admire the former with a higher degree of intensity than the latter.¹³⁸

It is common lore that affective episodes have a “what-it-is-likeness”, “phenomenology”, or “felt quality” and that affective dispositions, such as sentiments, lack such a felt quality, at least in themselves. We do not constantly feel something when we care for or love someone, for example. Rather, we have a tendency to experience certain kinds of affective episodes that target the object of our love or care depending on the circumstances.¹³⁹ This might become problematic if we take blame to be an affective disposition of some sort (as I do) as the most obvious way of accounting for intensity is by appeal to some phenomenological property.

I believe that we can overcome this worry. Tentatively, we can say that we privately blame agent *A* more intensely than another agent *B* when our private blame prompts affective episodes targeting *A* that are more intense than the affective episodes we have targeting *B*. For example, *A* privately blames *C* more than what *B* privately blames *D* when *A*'s blame sentiment prompts more intense feelings of resentment, disappointment, and such, than what *B*'s sentiment of blame does.

Some might want to insist that *duration* also determines the degree of an affective attitude. I do not think this is a relevant consideration. At least, I do not think it should be the *only* consideration. For the purposes of this chapter, I feel no pressing need to deny this very strongly. However, let me briefly state some reasons explaining why I am sceptical about accounting for degrees of affective attitudes by appealing to their duration.

When someone says they admire *Mona Lisa* more than *Girl with a Pearl Earring* they do not normally mean that they have admired the former for a longer period than the latter. Rather, what they usually mean is that they admire the former

¹³⁸ I assume, perhaps controversially, that we can make interpersonal comparisons. If *A* admires *Mona Lisa* very intensely and *B* admires it somewhat intensely, this means that *A* admires it more than *B*. I do not have in mind how each individual admires objects relative to other objects they admire. However, nothing important hinges on this, or on how, more generally, we make sense of intensity in detail. Rather, what is important for me is that we can (and reasonably so) appeal to intensity to make sense of attitudinal degrees.

¹³⁹ For more on this claim, see Deonna and Teroni (2011).

more intensely than the latter. Relatedly, to discover that, for ten days, I have admired *Mona Lisa* very intensely and *Girl with a Pearl Earring* somewhat less intensely is not to learn that I have admired them to the same degree. Rather, we would think that I admired the former more than the latter because I admired it more intensely.¹⁴⁰ Apart from this appeal to intuitions, we can note that appeal to duration would introduce theoretical difficulties. If I have sufficient reason to admire *Mona Lisa* to some extent, is admiring it for thirty days admiring it to some extent? Or should we use a percentage of expected lifetime? What happens after I have admired *Mona Lisa* for thirty days? Do I no longer have a sufficient reason to admire the painting? Would doing so be excessive? Must the thirty days be continuous? Given the intuitions and theoretical hurdles just mentioned, I will focus on the intensity of affective attitudes when accounting for their degrees.

In any case, the crucial point is that, by contrast, we do not account for the way actions come in degrees by appealing to their intensity (or duration). It makes little sense to speak of actions as having intensity (and note also that some actions lack meaningful duration: consider asking for an explanation or giving someone a fine). Rather, we appeal to something else when determining actions' degree of strength. Below, I consider two ways of accounting for degrees of actions that seem especially relevant given the purposes of this chapter.

Consider sending a mild letter of recommendation versus sending a strong letter of recommendation. The difference between mild and strong letters of recommendation is the intensity of attitude the letters convey: the latter conveys intense favouring, the former moderate favouring.¹⁴¹

It is a difficult question how we should understand in detail what it is to convey a certain intensity of an attitude through our overt actions. It is easy to adopt a naïve conception according to which the degree of intensity conveyed corresponds to different types of action. For example, one might believe that shouting

¹⁴⁰ It seems possible to (sincerely) say that one admires something greatly without having any attitude of admiration with a measurable degree of intensity at that particular time. In reply, consider what I just said about dispositions: granting that admiration is some kind of disposition, we make sense of its degree by appealing to the intensity of the emotions it prompts. If admiration is a disposition of some sort, one need not constantly feel something during the entire time one admires something.

¹⁴¹ The term "convey" is not perfect. It is associated with success: I convey my thoughts, or feelings, to you when you understand them. Nevertheless, hopefully, my terminology does not cause too much confusion.

necessarily conveys a more intense feeling of private blame than a normal conversational tone does. This is a mistake. It is, for instance, hard to overstate the immense intensity of hate that Johan, in Ingmar Bergman's *Scenes from a Marriage*, conveys when he quietly says "I hate you" to his wife. Relatedly, what qualifies as a normal conversational tone shifts wildly between different cultures and social contexts. Further complications are created by the fact that communication is rarely without friction. The intensity of attitude the target *registers* (or the audience registers) and the intensity of attitude the speaker *intends* to convey can come apart. For example, someone might have intended to send a strong letter of recommendation, and indeed thought that they had done so, even though the recipient read it as a mild letter of recommendation.

No matter in which way we choose to explain how the intensity of an attitude is conveyed through overt actions, for my purposes it suffices to note that we can convey varying degrees of intensity of our attitudes through our overt actions.

The other way of accounting for the degree of an overt action is by appealing to the harm the action causes. Harm admits of degrees: I can harm someone more or less. Likewise, it seems clear that overt blame can cause more or less harm – e.g., overt blame can result in a very painful feeling of guilt in the recipient or affect the recipient's interests negatively. On this way of accounting for the degree of overt blame, the degree to which you overtly blame someone is determined by how much the blameworthy agent is harmed by your overt blame.

There is a question about scope, and which harms should be considered relevant. Jeff McMahan's (2015) distinction between *wide* and *narrow* proportionality is helpful here. On the narrow scope reading it is only harms to the blameworthy agent that are to be considered. In contrast, on the wide scope reading we should also consider other harms, such as those suffered by individuals with close ties to the blameworthy agent. Here I will have the narrow scope reading in mind.

Another question concerns whether, in determining degrees of overt blame, we should be concerned with the *expected* harm of acts of overt blaming or their *actual* or *known* harm. More generally, the issue is whether we should be interested in *subjective proportionality* or *objective proportionality*. For the sake of simplicity, I will be concerned with the latter.¹⁴² It should be mentioned, however, that

¹⁴² See Tomlin (2018), in particular, on complexities that arise as soon as we want a satisfactory account of subjective proportionality in self-defence theory.

nothing important in my account of proportionate blame hinges on the adoption of an *ex-post* rather than *ex-ante* account of degrees of overt blame couched in terms of harm. In addition, with some added work my discussion in the next section can presumably be rephrased in the language of your preferred machinery of subjective proportionality without my main claims being affected.

A final question concerns whether, in determining degrees of overt blame, we should be concerned with the harm to the blameworthy agent or (what we might call) the badness of the harm to the blameworthy agent. Consider two blameworthy agents, *A* and *B*. *A* has already been harmed to a significant degree by overt blame but not *B*. Agent *C* overtly blames both *A* and *B*. Although *C* harms *A* and *B* equally by blaming them overtly, *C*'s overt blame of *B* seems worse than *C*'s overt blame of *A*.¹⁴³ I will be concerned with the badness of the harm, unless stated otherwise.

It is important to note that even if overt blame usually (though not necessarily) involves some harm being done to the blameworthy agent, overt blame need not be a form of punishment.¹⁴⁴ If you punish someone, "[...] then the harm that you cause is something that you intend to cause" (Zimmerman 2011a: 7-8). Merely intending to do something that you foresee will involve harm to someone is not to punish her. According to Michael McKenna, it is here that overt blame and punishment differ, because "the intention to harm is not *essential* to [overt] blame" (McKenna 2012: 145). Rather, the harm is in many cases merely a foreseen side-effect.

To the point: there can be cases of overt blame in which the blamer intends to cause harm to the blameworthy agent and thereby punish her – recall Chapter 2 (§2.8), where I explained that the blame sentiment can prompt punitive desires, which, when acted upon, can be a form of punishment. However, this is not true of all cases of overt blame. In other words, while some cases of (sincere) overt blaming are punitive, many are not.¹⁴⁵

¹⁴³ As noted by Tadros (2018: 26), this explanation seems to hinge on prioritarian intuitions about benefits mattering less the better off you are and burdens mattering more the worse off you are.

¹⁴⁴ For more on the claim that overt blame does not necessarily harm its target, see Brekke Carlsson (2017).

¹⁴⁵ Thank you, Björn Petersson, for inviting me to clarify this.

Having noted these differences between overt blame and punishment, I will argue in the next section that overt blame and punishment also seem to differ with respect to their proportionality conditions. But before I do this, I wish to highlight another difference, perhaps the most important one – between overt and private blame. The difference is that overt and private blame aggregate differently. More precisely, the intensity of private blame and the harm of overt blame do not aggregate in the same way. Even when a single instance of overt blame is not intended, and usually does not, to impose any morally relevant harm on the blameworthy agent, it can amount to a harming of the blameworthy agent *taken together with other instances of overt blame* in a way that is morally significant.¹⁴⁶ A thousand cuts can become a severe cut; being asked a thousand times “So, where are you from, *really?*” can constitute morally relevant harm, specifically by making you feel excluded from society or your community; a thousand instances of overt blame directed at you can constitute morally significant harm to you – e.g., they can make you feel very painful guilt, devastated, suicidal or depressed for months.

Importantly, affective attitudes do not aggregate in this way. Your feeling of indignation plus other agents’ feelings of indignation do not add up to a shared very intense feeling of indignation. This is a general feature of the way in which the intensity of attitudes aggregates. If we are at a somewhat funny stand-up show, each of us might be slightly amused. The intensity of our several attitudes does not add up so that, taken together, we are greatly amused. We are all, even taken together, only mildly amused. Our expressions of amusement via some overt action – our clapping our hands, for example – on the other hand, aggregate differently. Overt actions are in a sense “out in the world” in a way that differs from the way affective attitudes are.

¹⁴⁶It is unclear to me whether McKenna (2012) overlooks this fact in the following passage: “In short, the harm in blaming, even at its most extreme, is simply not nearly as severe as the harm that is possible in punishing, nor are the welfare interests that are threatened nearly as threatening to one’s overall well-being. For example, unlike blame, punishment might expose one to the possibility of a shortened life, absorbing physical pain, living in a less desirable social and physical environment, a minimised level of financial security, and so on” (McKenna 2012: 142).

5.2 Proportional Private Blame and Proportional Overt Blame

In this section, I provide the *beginnings* of an adequate account of proportional blame. I start with the idea most philosophers who have written about proportionate blame seem to share (see the introduction to this chapter). This is the idea that an instance of blame is proportionate if, and only if, its degree aligns suitably with how blameworthy the agent is – i.e., when blame “fits the transgression”. I shall consider whether this view is plausible in respect of private blame (§5.2.1), and then in respect of overt blame (§5.2.2). Before doing this, however, I wish to repeat a few words about how we make sense of degrees of blameworthiness.

As stated in Chapter 1 (§1.2.1), I grant that for an object x to be more (dis)valuable than some other object y is for it to be fitting to (dis)favour x more than y , and by “(dis)favour” I mean some pro- or con-attitude with a certain degree of intensity. Applied to blameworthiness: an agent A is more blameworthy than some other blameworthy agent B if, and only if, it is fitting to privately blame A more intensely than B .¹⁴⁷

As regards the question of what, exactly, makes one agent more blameworthy than another, we learned in Chapter 3 (§3.6.2 and §3.6.3) that the blameworthy agent’s knowledge (broadly construed) of the deontic status of her act or omission and the badness of her motive count as aggravating factors – i.e., factors that increase the intensity of private blame that it is fitting to direct upon her – and that stress, lack of good sleep, and such like, count as mitigating circumstances – i.e., factors that decrease the intensity of private blame that it is fitting to direct upon her.

¹⁴⁷ Our ascriptions of degrees of blameworthiness are not sharp. Neither is the intensity of our attitudes. Rather, when someone is, say, somewhat blameworthy a spectrum of intensity of blame will be considered proportionate. Others who have argued for this point include Andersson and Werkmäster (2020; 2022) and Tierney and Telech (2019).

5.2.1 Proportional Private Blame

Is it plausible to hold that an instance of privately blaming a blameworthy agent B for having φ -ed with intensity i is proportionate if, and only if, i matches the intensity i of private blame that it is fitting to direct upon B for having φ -ed (i.e., how blameworthy B is)? For example, does it sound reasonable to say that it is disproportionate to be furious with an agent (i.e., to entertain very intense private blame relative to her) who is merely slightly blameworthy? I believe the answer is “Yes”. Looking to other attitudes, we reach similar conclusions. An instance of admiring a certain object x with intensity i is proportionate if, and only if, i matches the intensity i of admiration that it is fitting to direct upon x , and so on. Therefore, unless there is a compelling reason *not* to understand proportionate private blame along the structural lines applying to other attitudes, I see no reason to reject the received view.

We can summarise the account of proportionate and disproportionate private blame as follows:

Proportionate private blame: an instance of privately blaming a blameworthy agent B for having φ -ed with intensity i is proportionate if, and only if, i matches the intensity i of private blame that it is fitting to direct upon B for having φ -ed.

Disproportionate private blame: an instance of privately blaming a blameworthy agent B for having φ -ed with intensity i is disproportionate if, and only if, i exceeds, or is less than, the intensity i of private blame that it is fitting to direct upon B for having φ -ed.

Given what I argued in Chapter 4, the question whether there is anything morally problematic about adopting disproportionate private blame arises. It is commonly thought that there is nothing morally problematic about privately admiring, loving or in some way disavowing something to an unfitting degree. And, as noted in Chapter 4, the fact that we cannot volitionally control our attitudes suggests that they cannot be wrong. I will not say more about this issue since, in this chapter, my aim is not to provide a complete account of proportional blame but rather the *beginnings* of a complete account of proportional blame.

5.2.2 Proportional Overt Blame

In §5.1, I said that, in trying to determine the degree of overt blame, we could refer either to the intensity of the private blame the overt blame conveyed or to the badness of the harm the overt blame caused to the blameworthy agent. From now, I will refer to the former way of making sense of overt blame's degree as $\text{degree}_{\text{conveyed}}$ and the latter way as $\text{degree}_{\text{harm}}$. I will first consider whether it is plausible to think that an act of overtly blaming a blameworthy agent B for having ϕ -ed is proportionate when its $\text{degree}_{\text{conveyed}}$ aligns suitably with how blameworthy B is for having ϕ -ed, and then whether it is reasonable to think that an act of overtly blaming a blameworthy agent B for having ϕ -ed is proportionate when its $\text{degree}_{\text{harm}}$ aligns suitably with how blameworthy B is for having ϕ -ed.

It sounds somewhat reasonable to say that an act of overtly blaming a blameworthy agent B is proportionate when the intensity of private blame conveyed through it matches the intensity of private blame that is fitting to direct upon B for having ϕ -ed. Further, it sounds fairly plausible to say that an act of overtly blaming a blameworthy agent B is disproportionate when the intensity of private blame conveyed exceeds, or is less than, the intensity of private blame that it is fitting to direct upon B for having ϕ -ed. It would be odd if the proportionality principles for private blame and for conveying private blame through overt blame were different, and there would be something insincere about conveying a degree of private blame one does not have or thinks it is proportionate to have. In addition, and generally speaking, we do think that it is something disproportionate about conveying a too strong, or mild, attitude via some relevant action – e.g., we believe that there is something disproportionate about conveying mild gratitude through an act of praising in response to a great favour. Thus, the common thought that blame should “fit the transgression” seems quite plausible here as well.

We can summarise the accounts of proportionate and disproportionate overt blame as follows:

Proportionate overt blame conveyed: an act of overtly blaming a blameworthy agent B for having ϕ -ed to $\text{degree}_{\text{conveyed}} n$ is proportionate if, and only if, n matches the intensity i of private blame that it is fitting to direct upon B for having ϕ -ed.

Disproportionate overt blame conveyed: an act of overtly blaming a blameworthy agent B for having ϕ -ed to degree_{conveyed} n is disproportionate if, and only if, n exceeds, or is less than, the intensity i of private blame that it is fitting to direct upon B for having ϕ -ed.

I believe that when we study the scenarios below, it becomes clear why the aforementioned proportionality principle for overt blame only manages to capture *a* sense of (dis)proportionate overt blame.

1. An agent, A , conveyed proportionate private blame through her overt blame to a blameworthy agent, B , but in a way that caused severe harm to B . In response, B takes responsibility for her action but complains that she has been blamed too much.
2. Several agents, A_1 - A_n , succeeded in conveying proportionate private blame through their overt blaming, and they did so in a way which, when considered in isolation, did not cause severe harm to the blameworthy agent, B . However, all of the overt blame, taken together, caused severe harm to B . In response to all of the overt blame, B takes responsibility for her action but complains that A_1 - A_n , taken together, have blamed her too much.
3. In a series of events, each agent A_1 - A_n succeeded in conveying proportionate private blame through their overt blaming. As a result of the series of overt blaming, B suffered some morally significant harm. She is clearly worse off after the overt blaming than she was before they were directed at her. However, she does not yet take responsibility for her blameworthy action. Agent C has committed a blameworthy act similar to B 's and has not taken responsibility for her action either. In contrast with B , C has not been blamed overtly for her blameworthy act and, therefore, she has not been harmed by anyone's overt blaming. Put differently, she is at the same level of wellbeing B was at before being blamed overtly by A_1 - A_n . You know of B and C 's blameworthy actions, and you know that they have not yet taken any responsibility for their actions. As a result, you overtly blame both of them. You conveyed proportionate private blame through your overt blaming, and your overt blaming harmed B and C equally (note that I am not, in this case, talking about the badness of the harm, see §5.1). In response to your overt

blaming, *B* and *C* takes responsibility for their actions. However, *B* complains that especially *your* overt blame was excessive. *C* does not make such a complaint.¹⁴⁸

Intuitively, in (1)-(3), *B* is right about the claim that she has been overtly blamed too much. This is the case even though all agents conveyed proportionate private blame. The overt blame in (1)-(3) is disproportionate, roughly, because *B* was harmed by it too much.

It is important to see that (1)-(3) need not show that *(Dis)proportionate overt blame conveyed* is completely unsound. As stated above, it seems to capture an important sense of (dis)proportionate overt blame. Focusing on (1), *B* could argue that the overt blame by *A* was proportionate in one sense – namely, in the sense that it conveyed proportionate private blame – but disproportionate in another sense – namely, in the sense that it harmed her too much.

Thus, (1)-(3) do not suggest that we should completely disregard the aforementioned proportionality principle for overt blame. Rather, they suggest that we need to amend it to make it clear that it only captures *a* sense of (dis)proportionate blame.¹⁴⁹ (1)-(3) also suggest that we need a proportionality principle for overt blame that captures the other sense in which *B* is overtly blamed too much in (1)-(3) – a principle that understands degrees of overt blame in terms of harm and explains why, when it does, the harm of overt blame becomes too much.

I now turn to consider whether the idea that blame is proportionate when it “fits the transgression” is plausible given that we make sense of overt blame’s degree by appealing to the badness of the harm it caused.¹⁵⁰ Below, I sketch a reason why we should reject such a view of (dis)proportionate overt blame.

Let us start by considering an analogy to the self-defence literature. Just considering the attacker’s degree of blameworthiness is insufficient to determine

¹⁴⁸ In (3) one can argue that the overall harm *B* suffers is disproportionate *and* that the final instance of overt blame, *C*’s, is disproportionate. The aim of (3) is to illustrate that the final instance of overt blame is disproportionate. In a moment, I will explain why the final instance of overt blame is disproportionate.

¹⁴⁹ In a moment, I will amend the aforementioned principle.

¹⁵⁰ Obviously, there is a question about whether we even can compare these variables in a rational way. Although an important question (and objection), I will not focus on it here.

whether a means of self-defence is proportionate or not. The badness involved in self-defence should be in proportion to the good realised by that action, and not just any good, but good in a narrow sense confined to the good of avoiding harm to self.¹⁵¹ If, in contrast, one were to say simply that a counterattack of the same degree is proportionate – regardless of whether any harm to self is avoided – one would no longer seem to be talking about proportionate self-defence, but rather about, at best, proportionate punishment.

Similar considerations apply to overt blame, in my view – if one were to say simply that overt blame of the same degree_{harm} is proportionate – regardless of whether (say) the blameworthy agent takes responsibility for her action in response– one would no longer seem to be talking about proportionate overt blame, but rather, at best, proportionate punishment. Thus, in order for an act of overt blame to be proportionate, the badness of the harm caused by blaming overtly must not exceed some specific good brought about by doing so.

In the case of self-defence, it seems analytically given that the specific good which the means of self-defence should be in proportion to is that of avoiding harm to self. In the case of overt blame, the specific good with which the degree_{harm} of overt blame should be in proportion to is not as clear. Tentatively, I believe that the good in question is connected to the nature, or function, of overt blame – e.g., standing up for the victim’s dignity or getting the blameworthy agent to sincerely acknowledge her wrongdoing and take responsibility for her act, or, perhaps, both.¹⁵² I will use the phrase “that the blameworthy agent takes responsibility for her action” as a placeholder for that good.

To summarise, we should not conflate overt blame with punishment by trying to weigh overt blame’s degree_{harm} with the blamed agent’s degree of blameworthiness for having ϕ -ed. Instead, taking inspiration from the debate on self-defence, we should weigh overt blame’s degree_{harm} against some specific good brought about by the act of overt blaming. I proposed “that the blameworthy agent takes responsibility for her action” as a place-holder for this specific good. In short: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed is disproportionate

¹⁵¹ For more on self-defence, just war and proportionality, see Frowe (2011) and McMahan (2014).

¹⁵² For more on the function of blame, see Wang (2021: 389) and §2.6.

when the overt blame's $\text{degree}_{\text{harm}}$ exceeds the goodness that is brought about by it, where the good in question is that B takes responsibility for her action.

Disproportionate overt blame harm: an act of overtly blaming a blameworthy agent B for having ϕ -ed to $\text{degree}_{\text{harm}} n$ is disproportionate if, and only if, n exceeds the goodness brought about by the overt blame in terms of B taking responsibility for her having ϕ -ed.

Proportionate overt blame harm: an act of overtly blaming a blameworthy agent B for having ϕ -ed to $\text{degree}_{\text{harm}} n$ is proportionate if, and only if, n does not exceed the goodness brought about by the overt blame in terms of B taking responsibility for her having ϕ -ed.¹⁵³

The blamer need not *intend* to get the blameworthy agent to take responsibility for her action when blaming her overtly. The blamer in question might very well intend to punish her. Whether a blamer's overt blame is proportionate is in this sense is not dependent on her intentions.

Further, the “act of overtly blaming” can refer either to a *single act* of overt blame, and the badness of the harm caused by it to B , or to a *collective act* of overt blame, and the overall badness of the harm caused by it to B . The term “collective act” is meant to be read in a broad way – i.e., it can refer to overt blame by a group agent, by a random group of unorganised overt blamers, or something in between.

We can rely on *Disproportionate overt blame harm* to explain (1)-(3). In (1), the badness of the harm A caused to B through her overt blaming exceeds the goodness A brought about by her overt blaming of B (i.e., B taking responsibility for her action). In (2), it is the badness of the harm caused by the collective overt blaming by A_1 - A_n that exceeds the brought about goodness of B taking responsibility for her action. Finally, in (3), your overt blaming of B is disproportionate, but your overt blaming of C is not, even though both conveyed the same proportionate private blame and caused the same amount of harm (not badness of the harm). One way of explaining (3) is as follows. The badness of the harm caused by A overtly blaming C is in proportion to the goodness thereby brought about (i.e., C 's acceptance of responsibility for her action), the badness

¹⁵³ I admit that it sounds odd to say that overt blame is *proportionate* when the harm it involved does not exceed the goodness it involved. It is more natural to say that it is not disproportionate. Relatedly, in cases where no good and no harm came out of your overt blame, it sounds odd to say that your overt blame is proportionate.

of the harm caused by *A* overtly blaming *B*, when *B* has already been overtly blamed, is disproportionate to the goodness brought about (i.e., *B*'s acceptance of her own responsibility for her action). Recall from §5.1, we seem to have the intuition that burdens matter more the worse off you are.

Before concluding this section, it is important to note that just as *(Dis)proportionate overt blame conveyed* was found to only capture *a* sense in which overt blame can be (dis)proportionate, the same is true of *(Dis)proportionate overt blame harm*. It only captures *a* sense in which overt blame can be (dis)proportionate – it does not, for example, capture the fact that overt blame can be disproportionate in the sense that it conveyed too intense private blame. In order to highlight that *(Dis)proportionate overt blame conveyed* and *(Dis)proportionate overt blame harm* each only capture *a* sense of proportionate overt blame; we can perhaps amend the principles by fixing what sense of “proportionate” they are concerned. For example, by subscripting (dis)proportionate_{conveyed} and (dis)proportionate_{harm}.

*Proportionate overt blame conveyed**: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed to degree_{conveyed} *n* is proportionate_{conveyed} if, and only if, *n* matches the intensity *i* of private blame that it is fitting to direct upon *B* for having ϕ -ed.

*Disproportionate overt blame conveyed**: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed to degree_{conveyed} *n* is disproportionate_{conveyed} if, and only if, *n* exceeds, or is less than, the intensity *i* of private blame that it is fitting to direct upon *B* for having ϕ -ed.

*Proportionate overt blame harm**: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed to degree_{harm} *n* is proportionate_{harm} if, and only if, *n* does not exceed the goodness brought about by the overt blame in terms of *B* taking responsibility for her having ϕ -ed.

*Disproportionate overt blame harm**: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed to degree_{harm} *n* is disproportionate_{harm} if, and only if, *n* exceeds the goodness brought about by the overt blame in terms of *B* taking responsibility for her having ϕ -ed.

The proportionality principles above seem reasonable. They should not to be viewed as competing principles, but as principles complementing each other.

5.3 Some Practical Implications

In this section, I briefly elaborate on some practical implications of my account of proportional blame.

I expect there to be plenty of real-life cases in which an agent's overt blame is proportionate in one sense but not in another, as the cases (1)–(3) discussed in the previous section illustrate. There is thus a question about which sense of proportional overt blame is more important in our current blaming practices.

While this is an interesting question, answering it is a task for future research. Intuitively, it is worse to harm someone too much via overt blame than it is to convey too strong an attitude of private blame through overt blame.

I also expect that there will be cases in which overt blame easily becomes disproportionate_{harm} given how the blameworthy agent is, as a person. Some blameworthy agents feel devastated upon learning that they have done something wrong. So, in these cases, depending on how we specify which good the degree_{harm} of overt blame should be in proportion to, it is likely that the badness of the harm caused by the overt blaming will exceed the goodness brought about by it.

Before concluding, I want to sketch one heuristic for ensuring that our overt blame is proportionate.

We should try to divide and spread the labour of overt blaming, especially in cases where a blameworthy agent has harmed or wronged several people. Rather than each and every agent (with standing to blame etc.) blaming the blameworthy agent overtly, it is advisable that they divide the labour of overt blaming and, for example, nominate someone to overtly blame her on behalf of the group. Not only is this less likely to lead to the imposition of disproportionate_{harm} overt blame (compare case (2) in section §5.2.2), but it is also likely to be more effective. Some people defend themselves aggressively instead of responding to the overt blame's demand when they feel that they are being blamed overtly by too many people.

In this chapter, I have focused on blame and proportionality. A responsible overt blamer probably also needs to respect something akin to the *necessity principle* in self-defence theory: roughly, the principle that if there is a plurality of possible overt blaming actions, then *ceteris paribus* you should choose the option that

minimises harm. Formulating and investigating a more precise necessity condition for overt blame is something that must be left for future research.

5.4 Conclusion

I believe this chapter has laid the ground for an approach to theorising about proportional blame that is fruitful. Of course, much more needs to be done. In particular, more will have to be said about the good that the badness of the harm of overt blaming should stand in proportion to, about how to draw the distinction between subjective and objective proportionality, about what is morally problematic about disproportionate blame – especially disproportionate private blame – and about how to divide the labour of overt blaming. Another question that I did not have time to discuss in this chapter is whether there is some sort of necessity principle for overt blame – i.e., the principle that if you can achieve the same good by blaming the blameworthy agent overtly in a less harmful way, it is impermissible for you to blame her overtly in the more harmful way. Finally, there is also a question whether the structure of the group consisting of overt blamers matters – i.e., whether it is a group agent or a random group of unorganised overt blamers, or something in between.

6 Testing the FA Analysis of Blameworthiness

In Chapter 1 (§1.2), I stated that in this thesis, I aim to formulate an informative FA analysis of blameworthiness and test it by appealing to some central claims made in the debate about the ethics of blaming. With insights from Chapters 2, 3, 4 and 5, the aim of this chapter is to do all of this.

In Chapter 1 (§1.2.1), I formulated the FA analysis of blameworthiness as follows:

FA blameworthiness: an agent A is blameworthy to degree n for having ϕ -ed if, and only if, (and because) it is fitting for anyone to blame A for having ϕ -ed to degree n .

“Fitting” denotes the presence of a “sufficient reason”. I left it unspecified what “blame” refers to.

As the reader might recall from §1.2, the FA analysis of value I depart from allows that we analyse value properties in terms of a fitting action, a fitting attitude, or some combination of fitting actions and attitudes. In addition, as the reader might recall from §1.2, the FA analysis of value I assume does not provide clear guidelines as regards what the attitudinal part of the analysis should refer to. The only guideline is that the attitude or action – or the combination of attitudes and actions – should in some relevant and direct way correspond to the value property we are concerned with.

Since both private and overt blame directly correspond to “blameworthy”, in this chapter, I will test an FA analysis of blameworthiness that take “blame” in the analysis to denote both private *and* overt blame. I will argue that such an analysis seems to entail unintuitive verdicts in cases involving standing-less blame and disproportional blame. In such cases, the analysis seems to yield the unintuitive result that agents who are clearly blameworthy are not blameworthy.

Consequently, I will argue that “blame” in the FA analysis of blameworthiness should refer to just private blame. I also provide some reasons for understanding the attitudinal part of the FA analysis of value strictly in terms of attitudes.

Here is the plan. In §6.1, I test the FA analysis of blameworthiness by appealing to some central claims made in the debate about the ethics of blaming. Then, in §6.2, I argue that we should restrict the FA analysis of value, including the FA analysis of blameworthiness, to analyse value properties by appealing to just fitting attitudes. By doing so, we avoid the challenges presented in §6.1.

6.1 FA Blameworthiness and the Ethics of Blame

In this section, I present a challenge to the FA analysis of blameworthiness. The challenge is that this analysis, in combination with central claims made in the debate about standing to blame and proportional blame, seems to yield the unintuitive result that agents who are clearly blameworthy are not blameworthy. I will also generalise this challenge to other value properties.

6.1.1 FA Blameworthiness and Standing to Blame

In Chapter 4, I stated that while norms of standing do not seem to apply to private blame at all, they do seem to apply to overt blame. While we cannot exercise any normative power by blaming someone privately, we can do so by blaming one another overtly; while the avoidance of meddling condition does not seem to apply to private blame at all, it does seem to apply to overt blame; and finally, while the conditions of standing cannot guide private blamers directly nor in any reasonable way, they can easily and reasonably guide overt blamers.

Now, suppose that Elliot has deliberately stolen a jar of candy from Mika and cannot provide a justification or excuse for his action. Elliot is blameworthy for the theft and it should therefore be fitting for anyone to blame him both privately and overtly (remember, I take “blame” in the FA analysis of blameworthiness to denote private and overt blame). However, it is not true that it is fitting for anyone to *overtly* blame Elliot for the theft. For Mika has previously deliberately stolen a jar of candy from Margot and has not made proper amends to her since then, and he cannot provide a justification or excuse for his action. Thus, Mika lacks

standing to blame Elliot overtly as a result of the hypocrisy this would involve and therefore, he lacks a sufficient reason to blame Elliot overtly for the theft (assuming that his *pro tanto* reason to not blame Elliot due to his lack of standing is not outweighed by his *pro tanto* reason(s) to blame him overtly).¹⁵⁴

If “blame” in the FA analysis of blameworthiness refers to private *and* overt blame, it seems that Elliot is not blameworthy for having stolen the jar of candy from Mika (since Mika lacks a sufficient reason to blame Elliot overtly, it is not true that it is fitting for anyone to blame Elliot privately and overtly). This is the wrong result.

6.1.2 FA Blameworthiness and Proportional Blame

In Chapter 5 (§5.2), I argued that different proportionality principles are associated with private and overt blame. An instance of privately blaming a blameworthy agent *B* for having φ -ed with intensity *i* is proportionate if, and only if, *i* matches the intensity *i* of private blame that it is fitting to direct upon *B* for having φ -ed. Overt blame, in contrast, is associated with two different proportionality principles. The first of these is parasitic on the proportionality condition for private blame: this is the principle that an act of overtly blaming a blameworthy agent *B* for having φ -ed to degree_{conveyed} *n* is proportionate_{conveyed} if, and only if, *n* matches the intensity *i* of private blame that it is fitting to direct upon *B* for having φ -ed. The second states that an act of overtly blaming a blameworthy agent *B* for having φ -ed to degree_{harm} *n* is proportionate_{harm} if, and only if, *n* does not exceed the goodness brought about by the overt blame in terms of *B* taking responsibility for having φ -ed.

Recall the case of Elliot, Mika and Margot (§6.1.1). Elliot is blameworthy for having stolen a jar of candy from Mika, and it should therefore be fitting for anyone to blame Elliot privately and overtly for the theft. But it is not true that it is fitting for anyone to blame Elliot privately and overtly for the theft – at least, not *overtly*. Margot lacks a sufficient reason to overtly blame Elliot for the theft

¹⁵⁴ One might immediately reply that Mika’s reason is of the wrong kind, roughly, as it does not determine Elliot’s blameworthiness. In §6.2.1, I will say more about this reply.

because doing so would be disproportionate_{harm}.¹⁵⁵ So, if “blame”, in the FA analysis of blameworthiness, referred to private *and* overt blame, it seems that Elliot would not be blameworthy for having stolen the jar of candy from Mika (since Margot lacks a sufficient reason to blame Elliot overtly, it is not true that it is fitting for anyone to blame Elliot privately and overtly). This is also the wrong result.

6.1.3 Summary and Generalising the Challenge

We sometimes have a sufficient reason not to overtly blame a clearly blameworthy agent because to do is excessive, or because we lack standing to blame her, and this does not show that the agent is not blameworthy. Therefore, the FA analysis of blameworthiness is extensionally inadequate and must therefore be false.

It seems that the problems raised in §6.1.1 and §6.1.2 can be generalised to other value properties. Concerning standing, we can imagine cases in which a certain agent is clearly praiseworthy but you lack standing to express praise to her because you do not know enough about her action or because you are not her peer.¹⁵⁶ So, you lack a sufficient reason to express your praise to her (assuming that your *pro tanto* reason to not express praise to her due to your lack of standing is not outweighed by your *pro tanto* reason(s) to praise her overtly). It is not true that it is fitting for anyone to praise her privately and overtly and, therefore, she is not praiseworthy (granting that we should analyse praiseworthiness in terms of fitting private and overt praise, or just overt praise).

Regarding proportionality, there comes a point at which others have already protected a valuable artefact and you lack a sufficient reason to protect it because to do so would be excessive, or pointless. Similarly, there comes a point at which, since others have already promoted, say, a book, you lack a sufficient reason to promote it yourself because doing so would be excessive, or pointless. However,

¹⁵⁵ I assume that Margot cannot overtly blame Elliot proportionately_{harm} – i.e., she cannot overtly blame Elliot without it being disproportionate_{harm} – and that she therefore lacks a sufficient reason to blame Elliot overtly.

¹⁵⁶ Whether there in fact are standing conditions for praise, or rather overt praise, is not something I will prove in this thesis. However, in the accompanying Appendix, I suggest that there might be norms of standing for overt praise. The possibility that there are conditions of standing for praise is enough for me to show that the arguments in this section can be generalised to all values that are subject to norms of standing.

just because you lack a sufficient reason to promote, or protect, these objects (it is not true that it is fitting for anyone to promote/protect them) does not show that these objects are not valuable (granting we analyse what is valuable in terms of it being fitting for anyone to protect/promote the object).

Thus, the challenges presented in this section need not be uniquely applicable to the FA analysis of blameworthiness. They may be challenges for FA analyses of other value properties as well.

6.2 The WKR Problem Again

As will be recalled from Chapter 1 (§1.2.2), the WKR problem suggests that the FA analysis of value is extensionally inadequate. In some cases, we have a sufficient reason to favour something that clearly has no value. In other cases, we have a sufficient reason not to favour something that clearly does have value. Therefore, the FA analysis of value must be false.

One might think that the problems I raised for the FA analysis of blameworthiness, and the FA analysis of other value properties, in §6.1 are mere variants of the WKR problem, and that whatever solution there is to it will also be applicable here.

I believe that the problems raised in §6.1 are variants of the WKR problem, in the sense that they suggest that the FA analysis of blameworthiness I am currently working with is extensionally inadequate – that blameworthiness and sufficient reasons to overtly blame can come apart – but that some solutions to the WKR problem are not at all helpful when we are seeking to solve the problems set out in §6.1. More precisely, while I do not believe it is very helpful to introduce the distinction between RKR and WKR to solve the problems raised in §6.1, I believe we can dissolve the tensions presented in §6.1.¹⁵⁷

¹⁵⁷ This is not the same thing as saying that I believe that we can dissolve the WKR problem more generally. I just think that we can dissolve this particular WKR problem to the FA analysis of blameworthiness.

6.2.1 The RKR and the WKR

As mentioned in Chapter 1 (§1.2.2), attempted explanations of how, exactly, to spell out the distinction between RKRs and WKRs in detail have now become a small cottage industry of their own. No consensus has so far emerged and all of the proposals in the literature face serious objections. I will show below that even if we grant that some of the proposed distinctions are sound, relying on those elucidated distinctions to solve the problems raised in §6.1 will not help. This is because they imply that the reasons pertaining to (say) standing are RKRs, not WKRs, and they need to be classified as WKRs in order to solve the challenges presented in §6.1.

In §1.2.2, I noted that it has been argued that there is a *motivational asymmetry* between RKRs and WKRs. The key thought here is that it is difficult, or even impossible, to be motivated by WKRs, but that this is not true of RKRs. It is difficult, or impossible, to admire a saucer of mud just because a demon has threatened one. Likewise, compare admiring da Vinci's *Mona Lisa* as a result of considering some properties of the painting – such as *Mona Lisa*'s smile – and admiring it as a result of considering a generous offer by a billionaire. It is not difficult to start admiring *Mona Lisa* as a consequence of studying the smile; it is difficult, if not impossible, to do so as a consequence of receiving a financial offer.

It does not seem difficult, or impossible, to refrain from blaming someone overtly because it would be meddlesome or hypocritical to do so. And it is not difficult to refrain from blaming someone overtly because the resultant blame is excessive. So, on this interpretation, reasons pertaining to lack of standing to blame someone overtly, or to disproportionality, have the hallmark of an RKR, not WKR.

In reply, one might argue that this claim about motivational asymmetry is meant to only apply to attitudes, as we can be moved to act for more or less whatever reason we take is sufficient. That might be true. However, according to another formulation of the RKR-WKR distinction, reasons pertaining to lack of standing to blame someone overtly, or to disproportionality, are also classified as RKRs.

According to the other suggestion (§1.2.2), RKRs are reasons shared by everyone engaged in a certain activity, because they are engaged in that activity. WKRs are not reasons which are shared by everyone engaged in a certain activity, because they are engaged in that activity. For example, everyone engaged in the activity of playing chess wants to win (let us suppose). Thus, for everyone engaged in that

activity, the fact that a certain move would help them win is a RKR to make that move, partly because they are engaged in that activity. An incentive offered by a demon to make a losing chess move does not give every other chess-player not so threatened a reason to make that move – such kinds of reasons are idiosyncratic.

Reasons appealing to standing and to proportionality seem to be reasons shared by everyone engaged in the activity of overt blaming, because they are engaged in the activity of overt blaming. They do not appear to be idiosyncratic. Everyone engaged in the activity of overt blaming, because they are engaged in the activity of overt blaming, refrain from blaming someone overtly when they lack standing to blame her overtly, and when it would be disproportional to do so (granting they have no stronger *pro tanto* reasons to blame her overtly nonetheless).

In sum, apart from the fact that no consensus has so far emerged on how to formulate the distinction between RKRs and WKRs, on some prominent accounts of the distinction, reasons pertaining to lack of standing to blame someone overtly or disproportionality seem to be wrongly characterised as RKRs to refrain from blaming someone overtly. Thus, the RKR and the WKR literature is not very helpful when we are trying to resolve the tension presented in §6.1 (because there we want to say that reasons pertaining to (say) standing are of the wrong kind).¹⁵⁸ If another alternative suggestion is capable of solving the problem raised in §6.1 in a plausible manner, it should be preferred. That said, it may still be true that the distinction between RKRs and WKRs that all will ultimately agree on will be applicable here and help us in stating that reasons pertaining to standing and proportionality are of the wrong kind.¹⁵⁹

¹⁵⁸ The proposal that RKRs are *object-given reasons* and that WKRs are *state-given reasons* seems to arrive at the right results, however. In short, object-given reasons “bear on the object” of that attitude, are provided by properties of the attitude’s object, and state-given reasons bear on the benefits, or costs, of being in the state of having that attitude, are provided by properties of the state of having the attitude. Considerations appealing to standing and proportionality do not bear on *x* but rather on the benefits, or costs, of being in the state of blaming *x*. In short, the reason why I have not focused on this distinction is that it does not seem to apply to actions such as overt blame. For more about this distinction, see Parfit (2011), Schroeder (2012), and Kieseewetter and Gertken (2016).

¹⁵⁹ Perhaps endorsing some circular distinction of the RKR and the WKR helps – roughly, RKRs are those that speak in favour of the object being valuable and WKRs are those that do not. Standing and proportionality do not speak for an agent being blameworthy, so they are the WKRs.

6.2.2 Dissolving the Tension

Another response to the WKR problem is to deny that it *is* a problem (see §1.2.2). Given the purposes of this chapter, I will need to provide a compelling explanation of why considerations of standing and proportionality do not cause problems for the FA analysis of blameworthiness. Fortunately, I can offer such an explanation, even if it is only a brief one.¹⁶⁰

It seems that what causes problems for the FA analysis of blameworthiness and the FA analysis of value in §6.1 is the (what we can call) conjunctive reading of the attitudinal part of the FA analysis of value – i.e., the idea that “(dis)favour” denotes attitudes *and* actions, and that “blame” refers to private *and* overt blame. The problems raised in §6.1 arose because it is possible to lack sufficient reason to blame or praise (etc.) someone overtly because there are issues of proportionality or standing. These issues do not afflict private attitudes, such as private blame, as there are no norms of standing for private blame and whether an instance of private blame is proportionate does not depend on what others do, or have done, or the blameworthy agent’s level of wellbeing.

Now I briefly show that there seems to be no compelling reason for including actions in the attitudinal part of the FA analysis of value, including the FA analysis of blameworthiness, and provide a reason for understanding the attitudinal part of the FA analysis of value, including the FA analysis of blameworthiness, strictly in terms of fitting attitudes. By doing this, we avoid the problems raised in §6.1 (since, as noted, they arose just because the attitudinal part included actions).

One reason for including actions in the attitudinal part of the FA analysis of value, including the FA analysis of blameworthiness, is, perhaps, this: there is a connection between value properties, such as “blameworthy”, and reasons for actions, such as reasons to overtly blame the blameworthy agent. The idea is that including actions in the attitudinal part of the FA analysis of value is a way to account for that connection.

Even if we grant that the attitudinal part of the FA analysis of value should strictly refer to attitudes, we can nonetheless explain the connection between values and reasons for actions. One possibility is to claim that while fitting attitudes are explanatorily prior to value, value is explanatorily prior to *pro tanto* reasons for

¹⁶⁰ Again, this is not the same as saying that we can dissolve the WKR problem more generally.

certain actions. This does not mean that all *pro tanto* reasons for actions are value-based. It means only that value is one potential source of *pro tanto* reasons for acting.¹⁶¹ On this line of explanation, André Aciman’s *Call Me by Your Name* is a good book because it is fitting for anyone to favour it (some attitude). Because it is a good book, there is a *pro tanto* reason to promote or recommend it. This *pro tanto* reason can be outweighed or undercut, for instance, if the person you recommend the book to has already read it. The case of overt blame without standing is another example. An agent is blameworthy because it is fitting for anyone to blame her privately, and because she is blameworthy, you have a *pro tanto* reason to blame her overtly. This *pro tanto* reason is undercut by your lack of standing which in turns explains why you lack a sufficient reason to blame her overtly. Consequently, it seems that we need not include actions in the attitudinal part of the FA analysis of value, including the FA analysis of blameworthiness, in order to make sense of the connection between values and reasons for actions.

Another reason for allowing that values be analysed in terms of fitting actions is the observation that there are some value properties where there is only a direct corresponding action, but no direct corresponding attitude. Consider, for instance, the value properties “dateable” (where that describes a person considered to be romantically or sexually desirable), “useful”, “punishable”, and “choiceworthy” and the following FA analyses of those properties:

FA dateable: An agent *A* is dateable if, and only if, (and because) it is fitting for anyone to date *A*.

FA useable: An object *x* is useable if, and only if, (and because) it is fitting for anyone to use *x*.

FA punishable: An agent *A* is punishable for having ϕ -ed if, and only if, (and because) it is fitting for anyone to punish *A* for having ϕ -ed.

FA choiceworthy: A course of action ϕ is choiceworthy if, and only if, (and because) it is fitting for anyone to choose ϕ .

¹⁶¹ This idea is similar to Skorupski’s bridge principle: whatever facts give agent *A* reason to feel, say, blame give *A* *pro tanto* reason to do the blame-prompted action, such as a *pro tanto* reason to demand an apology from the wrongdoer, *in virtue of being a reason to feel blame*. For more about this principle, see Skorupski (2010a).

The verbs “use”, “date”, “punish” and “choose” refer to actions, and there is no immediate corresponding attitude-version of, for example, “choose”. If we restrict the FA analysis of value to fitting attitudes, it will fail to make sense of a range of value properties, or so the argument goes.

I believe this objection is interesting and challenging, but we can overcome it.

The fact that the value properties “dateable” and “blameworthy” directly correspond to “date” and “overt blame” do not entail that “dateable” and “blameworthy” must be analysed in terms of it being fitting to date or overtly blame the agent. It seems we can just as easily analyse them strictly in terms of attitudes: for example, it is fitting to *desire* to date the agent (an attitude that also corresponds to the value property “dateable”, although not directly) or it being fitting to privately blame the agent (an attitude that also corresponds directly to the value property “blameworthy”).

Likewise, several attitudes and actions correspond directly to the good – e.g., “like”, “favour”, “protect”, “desire” and “promote”. However, it does not seem odd to say that an object’s being good should be analysed just in terms of it being fitting to, say, favour it.

To the point, it seems to me that we can restrict the FA analysis of value to fitting attitudes without losing any explanatory power – such as being able to explain the connection between values and reasons for actions – or failing to make sense of certain value properties – such as “choiceworthy”. In addition, by restricting the FA analysis of value to fitting attitudes, we also end up with a more elegant analysis and an analysis that avoids the challenges presented in §6.1.

In reply, one might perhaps argue that we should distinguish between *attitude-oriented values* (i.e., values that correspond directly to just attitudes or also to attitudes) and *action-oriented values* (i.e., values that correspond directly to *just* actions), and treat them differently. For example, as the reader might remember from Chapter 1 (§1.2), actions do not come in degrees in a way that is helpful for FA scholars that want to account for values’ degree (unlike attitudes, actions are not more or less intense). Therefore, on this reply, we should make sense of action-oriented values’ degree, such as choiceworthiness’ degree, by appealing to how fitting it is to choose the object, not in terms of how much choosing is fitting, and

attitude-oriented values, such as “blameworthiness”, by appealing to how much private blame is fitting.¹⁶²

Likewise, on this reply, I am correct in that we should analyse blameworthiness – which is an attitude-oriented value since it does not directly correspond to just an action (“overt blame”) but also to an attitude (“private blame”) – just in terms of private blame but incorrect in that we should try to analyse action-oriented values – such as “useable” that corresponds directly to just the action “use” – in terms of fitting attitudes. In short, attitude-oriented values should be analysed just in terms of fitting attitudes and action-oriented values just in terms of fitting actions.

Future research can evaluate this reply. Tentatively, I am sympathetic to it. By accepting it, we avoid the challenges to the FA analysis of blameworthiness mentioned in §6.1 (we also probably avoid the challenges to the FA analysis of praiseworthiness and goodness by accepting it).¹⁶³ We also get an elegant analysis of blameworthiness. And we can continue to analyse value properties that directly correspond to just an action in terms of a fitting action, instead of having to reformulate existing FA analyses of those values.

6.3 Conclusion

In Chapter 1 (§1.2.2), I said that to make the FA analysis of blameworthiness more informative, we need to say more about what is meant by “blame” and what makes it fitting for anyone to blame an agent *A* for having ϕ -ed to degree *n*. I now have answers to those questions.

In Chapter 3, I defended the view that what makes it fitting for anyone to blame agent *A* for having ϕ -ed is the fact that *A* ϕ -ed deliberately despite being aware of the fact(s) that constitutes a normative reason(s) which, taken together, makes it wrong for *A* to ϕ , and *A* could and should have believed that ϕ -ing is wrong. And in this chapter, I have briefly argued that we should restrict the FA analysis of value – or rather, the FA analysis of *attitude-oriented values* – including the FA analysis of FA blameworthiness, so that we analyse attitude-oriented values by

¹⁶² Thank you, Wlodek Rabinowicz for mentioning this potential distinction to me.

¹⁶³ In Appendix, I suggest that there is private praise (meaning some attitude).

appealing to just fitting attitudes, such as private blame. By “private blame” I mean, as will be recalled from Chapter 2 (§2.8-§2.9), the mere harbouring of the sentiment of blame. Consequently, I believe we should reformulate the FA analysis of blameworthiness as follows:

*FA blameworthiness**: an agent A is blameworthy to degree n for having ϕ -ed if, and only if, (and because) it is fitting for anyone to privately blame A for having ϕ -ed to degree n .

7 Summary and Critical Discussion

At the beginning of this thesis, I presented five questions that I would be aiming to answer:

1. What is the nature of blame?
2. What makes an agent blameworthy for having performed a particular action or omission?
3. Do norms of standing to blame apply to private blame?
4. What is it for an instance of blame to be proportionate?
5. How do the ideas of standing to blame and proportional blame relate to the fitting attitude analysis of blameworthiness? For example, are they in tension?

In this concluding chapter, I provide a summary of my answers to these questions. I also highlight my contribution to the debate, exercise some self-criticism about what I think I have not managed to show, and sketch directions for future research. I will take topics in the same order as the chapters in this monograph.

In Chapter 2, I argued that blame is not to be identified solely with a judgment. It is one thing to judge an agent blameworthy, another to blame her. Blame is not to be identified solely with an overt act either. The main reason for this is that we can blame “privately”, without performing any overt action. Nor is blame to be identified solely with an angry emotion, since it is not always something involving anger. Nor, again, is blame to be identified simply with a revision of expectations. There is more to blame than that – e.g., we also tend to experience various emotions. Blame is not a matter of believing that a certain agent acted badly and desiring that she acted differently. One can believe and desire in that way without blaming someone. Finally, blame is not to be identified with a function, or with a process. In Chapter 2, I argued that blame should instead be characterized as a sentiment. More precisely, it should be defined as a disposition that manifests

itself in the form of various emotions, beliefs and desires to act in a range of different circumstances.

As I stated in (§2.3), I am sceptical about the distinction between moral and non-moral blame, and indeed about the distinctions between epistemic blame, aesthetic blame, and so forth. I claimed there that it sounded odd to say that someone merited blame for having made (say) a mathematical mistake. It was not my aim to provide a complete refutation of these distinctions. That is a task for future research. It is important to note, however, that the soundness of the above distinctions need not cause any problems for my account of the nature of blame (§2.8).

I did not provide much detail, in Chapter 2, on the manifestations and triggering conditions of the sentiment of blame. Future research can provide an adequate list of such manifestations and triggering conditions. This task might prove to be difficult, however. Multi-track dispositions have several triggering conditions and manifestations, T1 + M1, T2 + M2, and so forth. There is a risk that it will be too hard – or even impossible – to provide an adequate list of the manifestations and triggering conditions.

Finally, more can be said about the nature of dispositions. In Chapter 2, I assumed that dispositions are powers. This view is not shared by everyone. For instance, it is common to opt for some version of the conditional analysis of dispositions:

Conditional analysis of dispositions: if x were C, x would M.

As is well known, this analysis has attracted a lot of criticism. It is claimed that it is vulnerable to counter-examples, such as ones involving finks and masks, and the like. It is also claimed that it is structurally inadequate and not being able to make sense of degrees, among others things.¹⁶⁴ Future research can compare these conceptions of dispositions in order to determine which we should adopt in our analysis of the nature of blame.

My main contributions to the debate about the nature of blame is a novel account of blame that I hope avoids important objections to rival accounts, while doing justice to many of our fundamental intuitions about the nature of blame. The account I have developed does not imply that blame is obviously a very bad or harmful attitude. This is a significant upshot. If blame is not that, it is not

¹⁶⁴ See, for example, Manley and Wasserman (2008), and Vetter (2015).

something that clearly needs to be abandoned or that we need very strong reasons to adopt. Finally, I have shown that we need not be sceptical about finding a unified and informative analysis of the nature of blame.

Let me move on to the question of what makes agents blameworthy for their actions or omissions. Surveying competing accounts, I argued in Chapter 3 that agents are not blameworthy in virtue of willingly having performed actions or omissions despite the belief that these were wrong. Nor are they blameworthy in virtue of deliberately having performed actions or omissions for the motive that make them wrong to perform. Rather, I argued agents are blameworthy for performing particular actions or omissions in virtue of their having done so despite being aware of the facts that constitute normative reasons which, taken together, make their actions or omissions wrong, and they could and should have believed that their act or omission was wrong.

In Chapter 3, I made some crucial assumptions. I assumed subjectivism about wrongness, where subjectivism is understood to track the facts of which agents are aware of. I also assumed that the recklessness objection is sufficient to refute objectivism. There are answers to this objection that I have not evaluated. And there are other subjectivist accounts of wrongness that I did not consider in the chapter, such as Elinor Mason's (2019). Roughly, Mason understands subjectivism about rightness and wrongness in terms of trying to do well by what she calls Morality. Perhaps some other understanding is preferable to the one I assumed. Future research can evaluate these answers and competing accounts.

Further, the wrong-making view*** has "capacitarian" elements because it adds a condition stating that the agent could and should have believed that her action or omission was wrong. Several capacitarian accounts have been proposed in the past.¹⁶⁵ Future research can compare the wrong-making view*** with competing accounts and highlight the similarities and differences between them. It can assess as well which capacitarian account is the most plausible.

Furthermore, in Chapter 3, I did not evaluate views according to which agents are blameworthy to the extent their actions or omissions reveal their poor quality of will, where "poor quality of will" is understood in terms of bad cares or morally problematic evaluative judgments. Future research can compare my view of

¹⁶⁵ Compare, for instance, Ayars (2021), Clarke (2017), Murray (2017) and Rudy-Hiller (2017).

blameworthiness with that view – whether they, for example, are in tension or are just concerned with different notions of blameworthiness.

Finally, in that chapter, more needs to be said about the challenges facing the wrong-making view**, such as the challenge that we can be blameworthy for right, or neutral, actions or omissions, not just wrong actions or omissions.

In sum, my main contribution to the debate over what makes an agent blameworthy is this: I have provided detailed presentations of two leading accounts of blameworthiness; and I have developed a tentative and novel theory of what makes agents blameworthy for having performed particular actions or omissions that allows us to make sense of important intuitions we have about blameworthiness.

The ethics of blame sets out conditions determining whether an instance of blaming is permissible. Nearly everyone agrees that two such conditions are that blame should be proportionate and that the blamer should have standing to blame the blameworthy agent. In Chapters 4 and 5, I argued that these two conditions are vague. More precisely, I argued that key features of the idea of standing to blame do not apply to private blame, and that as a consequence we are not justified in holding that there are norms of standing to blame privately. Further, I argued that private and overt blame are associated with different proportionality principles.

More could have been said in these chapters. More needs to be said about what sort of control we have over private attitudes, including private blame. In Chapter 4, I assumed that we cannot just give up our private blame on the basis of any reason, such as a reason pertaining to standing to blame.

Further, in Chapter 5, I was silent on the good that overt blame should be proportionate_{harm} to. Future research can answer this question.

I did formulate objective proportionality principles for overt blame. However, it would be interesting to carve out a subjective proportionality principle for overt blame. And it would also be interesting to investigate whether a necessity principle of some kind applies to overt blaming – e.g., the principle that when a less harmful alternative achieves the same goal as your overt blaming, it is impermissible for you to blame someone overtly.

My main contribution to the debate about standing to blame and proportional blame is as follows. I have defended scepticism about norms of standing pertaining to private blame. I have detailed significant differences in the character of overt and private blame, and I have made proposals as to how to state proportionality principles for overt and private blame, respectively. The general lesson to take home from this is that any satisfactory account of the ethics of blame needs to pay close attention to the distinction between private and overt blame.

In Chapter 6, I argued that the claim (i) you lack sufficient reason to blame a clearly blameworthy agent when you lack standing to blame her, and the claim (ii) that you lack sufficient reason to blame such an agent when it would be disproportional to do so, create problems for defenders of the FA analysis of blameworthiness. In such cases, this analysis seems to imply that clearly blameworthy agents are not blameworthy. I briefly argued that we avoid these problems if we restrict the FA analysis of blameworthiness, and the FA analysis of value more generally, so that they appeal only to fitting private attitudes such as fitting private blame.

Although I considered several replies to the challenges from standing and proportionality, I did not consider the reply that we should opt for a complex FA analysis of value. Such an analysis, focusing on blameworthiness, could look like this: an agent *A* is blameworthy to degree *n* for having ϕ -ed if, and only if, (and because) it is fitting for anyone to privately blame *A* for having ϕ -ed and fitting for anyone to overtly blame *A* for having ϕ -ed (unless one lacks standing to blame *A*, etc.). Future research can investigate this reply. Tentatively, I believe that we should opt for this reply if we *must* analyse, say, blameworthiness in terms of fitting private *and* overt blame, and if the other reply I considered in Chapter 6 is not plausible – e.g., the reply that we should distinguish between action-oriented values and attitude-oriented values and treat them differently.

Throughout Chapter 1 and Chapter 6, I assumed that formulating and testing the FA analysis of blameworthiness was a viable undertaking. Some may think the pursuit of such an analysis does not make sense. According to Rach Cosker-Rowland (2019), for example, we will struggle to provide a non-circular analysis of moral blameworthiness because it is hard to distinguish between non-moral and moral blame in a non-circular fashion. As I said in Chapter 2 (§2.2), I too am sceptical about the feasibility of making a distinction between moral and non-moral blame (although for other reasons). However, at least initially, it seems that

my FA analysis of blameworthiness is not circular, as the content of the sentiment of blame is not “*A* is blameworthy”.

My main contribution to the debate about the FA analysis of value is an informative FA analysis of blameworthiness, a novel distinction between action-oriented values and attitude-oriented values, and the beginnings of a discussion about what the attitudinal part of the FA analysis of value should refer to.

As will perhaps be evident after reading this thesis, current philosophical discussion of blame is still very much in its infancy. I hope this thesis has helped to structure our thoughts about blame in a fruitful manner. I also hope it has pointed to research gaps that need to be filled and even filled some of those gaps.

8 Appendix: On Praise

A widely (though certainly not universally) held view among both lay people and philosophers is that praise and blame are, more or less, symmetrical. In other words, the view is that similar factors are relevant for determining blameworthiness and praiseworthiness, or that the nature of praise and blame are more or less two sides of the same coin, or something to that effect. It might therefore be considered to be an advantage if my account of blame easily translates to praise. In this Appendix, I will *briefly* investigate whether we can generalise my conclusions about blame to praise. I will argue that we can fairly easily generalise some conclusions but not all of them.

8.1 On the Nature of Praise

In Chapter 2 (§2.1), I presented some core intuitions we have concerning blame's nature. These are that blame is:

- a) about something;
- b) connected with various emotions: e.g., anger, disappointment, sadness, guilt;
- c) connected with motivations and desires to act or actions: e.g., motivation to demand an apology, desire to demand an explanation, social distancing, ostracism;
- d) connected with thoughts and evaluations: e.g., the evaluation that a particular agent ought to apologise, the thought that it is good that others blame a particular agent; and
- e) relatively long lasting, lasting for several minutes, hours, days or even years, usually until we feel proper amends have been made.

Do we have similar intuitions about praise? I think we do. We think that praise is about something. Further, we think that praise is connected with emotions, albeit perhaps not as many emotions as blame. Positive emotions such as “admiration”, “approbation”, and “gratitude” come to my mind. Further, we think that praise is robustly connected with certain actions and kinds of desire to act. Generally, we desire to tell the praiseworthy agent, or others, about her praiseworthy deed, celebrate her on the basis of her act, encourage others to collaborate with her, and such like. And we think that praise is connected to certain characteristic thoughts, or evaluations, such as the thought that it is good that others praise the praiseworthy.

What about the intuition that praise is relatively long-lasting? There is a clear end to blame, but not so much to praise. We usually blame someone until she has apologised properly. This often takes some time. In addition, blame usually ends with forgiveness – an attitude adopted by the blamer, often following an apology or the making of amends. We cannot say something similar about praise. For one thing, there seems to be no positive analogue to “apologise”. Would it be for the person being praised to say “thank you”? We do not seem to stop praising someone after having received a “thank you” in acknowledgment of our “kind words”, for example. And there is no analogue to “forgiveness” in the case of praise. That said, I believe praise is relatively long-lasting. It seems to me that once we find someone praiseworthy for something, and start praising her, there is no end to our praise. Whenever we are reminded about the praiseworthy agent and her conduct, we are prone to praise her for her actions.

In one of the preliminary sections of Chapter 2 (§2.3), I said that it is common to distinguish between moral blame, non-moral blame, epistemic blame, self-blame, other-blame, and so forth. Is it common to distinguish between the various corresponding forms of praise as well? I believe we do ordinarily distinguish between various forms of praise – e.g., between moral praise and non-moral praise, and so forth. For example, I praise my son for having learned how to count to ten, and this kind of praise is not moral, and one can praise both others and oneself.

When analysing specific accounts of the nature of blame, I presented some objections, and these objections pointed to important distinctions and phenomena that we want any adequate account of blame to account for. To recall two of these: first, it is one thing to judge an agent blameworthy and another thing to blame her (§2.4.1); and second, we believe that we can blame an agent without

performing any overt action and without letting anyone know about our blame – i.e., we can blame someone “privately” (§2.4.2). Do we need to account for similar phenomena when formulating a theory of praise? I think we do.

It does not sound misguided to claim that it is one thing to judge an agent to be praiseworthy and another thing to praise her. Consider the following case provided by Daniel Telech (2022):

Satan and the amoralist can presumably judge Alica to have acted praiseworthy, in the sense of taking her to have done ‘the right thing for the right reasons’, say – without thereby praising her. In reflecting on Alica’s having placed another’s interests before her own, Satan might think something like: ‘she benevolently did the morally right thing – *what a sucker!*’ Presumably, this is not praise. A general way to put the problem is that the judgment view cannot obviously accommodate the way in which praiser are favorably disposed toward those they praise. (Telech 2022: 3)

I share Telech’s feeling that there is something missing from an account of praise that equates it solely with the making of a certain judgment. As with the judgment account of the nature of blame, what is missing is some reactive or emotive component. We tend to react in certain ways towards those we praise. We are prone to feel admiration, joy or gratitude for what they have done, and we desire to collaborate with or celebrate her in some way.

It also seems that, just as one can blame someone privately, “in one’s heart”, one can praise someone privately. Consider the following case:

Aiding: Sara mocks David. Ashorina sees this and intervenes. She does so because she wants to stand up for David’s dignity. Ashorina’s action is clearly praiseworthy. I quietly observe the scene. I do not dare to praise Ashorina overtly because I do not want Sara to mock me afterwards. However, in my heart, I feel gratitude towards Ashorina and a desire to celebrate her, be near her, as a result of the action she performed.

In *Aiding*, I praise Ashorina even though I do not perform any overt action. This suggests that praise cannot be equated with the performance of an overt act.

So, it seems that people tend to have similar intuitions about the nature of praise and the nature of blame, and we also think that we have to account for some similar phenomena and objections regarding both. In Chapter 2 (§2.8), I argued

that understanding blame as a sentiment helps us to make sense of all our intuitions about the nature of blame, as well as making it possible for us to understand important phenomena and objections.

This is not the place to defend and develop a sentiment-based account of praise. However, I do not think it sounds completely unreasonable to venture the thought that when we praise someone, we are, in various circumstances, prone to experience a range of emotions, thoughts and desires to act with respect to her on the basis of the action or omission she has performed. Characteristic emotions, actions and desires of this sort include, but are not limited to, the following.

In other-praising:

Characteristic emotions: gratitude, joy, admiration, approbation.

Characteristic thoughts: believing that it is fitting for anyone to praise the agent, that it is good when other agents praise the agent, that it is good to collaborate with the agent.

Characteristic desires: wanting to invite the agent to share our joy in her action, to recommend others to collaborate with the agent, to help the agent in turn.

In self-praising:

Characteristic emotions: pride, joy, pleasure.

Characteristic thoughts: believing that one has done something good, that it is fitting for others to praise one, that one has earned some benefit.

Characteristic desires: wanting to benefit oneself.

To sum up, the claim that blame and praise are more or less two sides of the same coin has at least some *prima facie* appeal.

8.2 On What Makes Agents Praiseworthy

I began Chapter 3 by presenting the distinction between *de re* and *de dicto* concerns. Roughly, the former concern is about whether agents are motivated by right-making or wrong-making features and the latter about whether they are

motivated by what they believe is right or wrong. In the chapter, I evaluated two accounts of what makes agents blameworthy for having performed particular actions or omissions: the witting wrongdoing view and the wrong-making view. The first of these can be interpreted as a *de dicto* view, as it states that agents are blameworthy for actions or omissions if, and only if, they willingly performed these actions or omissions despite consciously believing that it is wrong to do so. The second can be interpreted as a *de re* view, since it states that agents are blameworthy for performing particular actions or omissions if, and only if, they performed these actions or omissions deliberately and their motivating reason(s) coincides with the normative reason(s) that makes the relevant actions or omissions wrong.

The main contenders in the debate about positive moral worth, or praiseworthiness, are what we might call the *rightness itself view* and the *rightness-making view*.¹⁶⁶ According to the former, roughly, an agent is praiseworthy for performing a right action or omission ϕ if, and only if, she ϕ -ed with the motive that ϕ -ing is right.¹⁶⁷ As is perhaps evident, the witting wrongdoing view echoes the rightness itself view.

According to the rightness-making view, by contrast, and crudely, an agent is praiseworthy for performing a right action or omission ϕ if, and only if, her motivating reason(s) coincides with the normative reason(s) that makes performing said action or omission right.¹⁶⁸ As is clear, the wrong-making view echoes the rightness-making view.

¹⁶⁶ Many scholars writing on moral worth and praiseworthiness use these notions interchangeably. It is beyond the scope of this thesis to consider whether the notions are different.

¹⁶⁷ The details of the rightness itself view are contested. Sliwa, a defender of a version of the view, formulates her view as follows: “a morally right action has moral worth if and only if it is motivated by concern for doing what’s right (conative requirement) and by knowledge that it is the right thing to do (knowledge requirement)” (Sliwa 2016: 394). In contrast, Johnson King, another rightness itself theorist, formulates her view like this: “an act has moral worth just in case it is an instance of someone’s *deliberately* doing the right thing” (Johnson King 2020: 201). The main difference between Sliwa and Johnson King’s views is that Johnson King does not require that the agent should *know* that she is doing the right thing. Instead, she thinks that only minimal foresight that one is in fact doing right is required for moral worth.

¹⁶⁸ As with the rightness itself view, there are different formulations of the right-making view. For instance, Markovits (2010) presents it as follows: an action is morally worthy if and only if – and to the degree that – the noninstrumental reasons motivating the action coincide with the noninstrumental reasons that morally justify its performance (Markovits 2010: 230 & 238).

Each view holds that only one kind of motivation is relevant to the moral worth of actions or omissions. In other words, the motivations that one view represents as necessary and sufficient for praiseworthiness are the very ones that the other view represents as irrelevant.

In Chapter 3 (§3.3), I also presented two cases that any reasonable account of what makes agents blameworthy needs to account for adequately: first, agents who do what is good, or the right thing, despite believing that what they are doing is wrong, or bad, are not blameworthy, and second, agents who do the bad or wrong thing despite believing that it is right, or good, are blameworthy. In §3.3, I also claimed that a good account of what makes agents blameworthy needs to make sense of these intuitions in an adequate manner.

Now, do we believe that a plausible account of what makes agents praiseworthy needs to account for the contrastive cases? I believe that it is at this point that it starts to become more conspicuous that blameworthiness and praiseworthiness are asymmetric at least to some extent. It seems to me that we do not think that we need to account for similar contrastive cases when we construct an adequate account of what makes agents praiseworthy for their actions or omissions.

The case contrasting with that of an agent who does something right, or good, while believing she is doing something bad, or wrong – i.e., Huck Finn – is one in which the agent does something wrong, or bad, while believing she is doing something right, or good – e.g., Andy. Intuitively, agents such as Andy are not praiseworthy at all for their deeds. For example, there seems to be nothing praiseworthy at all about an SS Officer who sends hundreds of Jews, Romans or political activists to deadly concentration camps while believing that what he is doing is right or good. Thus, we have results that mirror one another: just as we think that agents such as Huck are not blameworthy, we think that agents such as Andy are not praiseworthy.

The case contrasting with that of an agent who does something wrong, or bad, while believing that she is doing what is right, or good, (e.g., Andy) is one in which the agent does something right, or good, while believing that she is doing something wrong, or bad – as happened in the scenarios involving Huck and Pam.

It is less clear whether agents who do what is right, or good, while believing they are doing something wrong, or bad, are praiseworthy. There is a lively debate whether Huck is praiseworthy for helping Jim escape while believing that it is

wrong to help Jim escape. Intuitions about whether agents such as Huck are praiseworthy are divided. However, the intuitions about whether agents such as Andy are blameworthy are less divided – as the reader might recall from Chapter 3, agents like Andy seem to be the prime examples of blameworthy agents.

In sum, we believe that agents that do wrong, or bad, while believing they do right, or good, are not praiseworthy and we are undecided whether agents that do right, or good, but believe they do wrong, or bad, are praiseworthy. This suggests that we need to account for intuitions that do not mirror those we have considered in connection with blameworthiness.

In Chapter 3, I argued that neither the witting wrongdoing view nor the wrong-making view – i.e., neither of the views about blameworthiness that mirrored the rightness itself view or the right-making view – were plausible. They could not account for the cases about Andy, Pam and Huck adequately. So, if the right-making view or the rightness itself view turn out to be plausible, or clearly correct, we have a further asymmetry between blameworthiness and praiseworthiness.

At the end of Chapter 3, I started by defending a modification of the wrong-making view. This modified view states that an agent A is blameworthy for an action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) ϕ -ed despite being aware of the fact(s) that constitutes a normative reason(s) which, taken together, makes it wrong for A to ϕ . Is it reasonable to think that a contrastive view is plausible when it comes to praiseworthiness? That is, can we say that an agent A is praiseworthy for an action or omission ϕ if, and only if, (i) A had volitional control over ϕ , and (ii) ϕ -ed while being aware of the fact(s) that constitutes a normative reason(s) which, taken together, makes it right for A to ϕ ? It is not obvious that the answer is “Yes”.

When I contemplate praiseworthy agents, I think of agents who have performed (high stakes) *supererogatory* actions, not merely right actions: for example, those who, at with great costs to themselves, walk into burning buildings to save innocent people.

Perhaps “right” is not the opposite of “wrong”. Perhaps it is more reasonable to think that “supererogation” is the diametrical opposite of “wrong”: so, wrong actions are merely “at odds with under the call of duty” while supererogatory actions go “beyond the call of duty”. If this is correct, we can rephrase the view above as follows: an agent A is praiseworthy for having performed an action or

omission φ if, and only if, (i) A had volitional control over φ , and (ii) φ -ed while being aware of the fact(s) that constitutes a normative reason(s) which, taken together, makes it supererogatory for A to φ .

Now to a more serious worry. It seems that defenders of the view I sketched above will have problems satisfying *the accidentality constraint* – or at least will have trouble doing so on some formulations of the constraint. This is a constraint that all of those involved in the debate over positive moral worth and praiseworthiness take to be one that an adequate account of praiseworthiness or positive moral worth must explain.

There are different suggestions as to what it is for an act to not be accidental. Some allude to some kind of modal or counterfactual stability or robustness. For example, Sliwa argues that:

The thought is that morally worthy actions are motivated in a way that makes their rightness neither “contingent” nor “precarious” – they are counterfactually robust. (Sliwa 2016: 394)

That is, roughly, the action is non-accidental if, and only if, the agent’s motive would lead her to perform the right action in all or most nearby possible worlds. If the same motivation in similar counterfactual scenarios would lead her to perform a neutral or wrong action, it was just an accident that she did the right thing.¹⁶⁹

Others argue that non-accidentality has less to do with modal robustness. It is centrally a matter of creditworthiness- or praiseworthiness. For example, Keshav Singh writes:

The sense of non-accidentality we’re interested in here is the sense in which the appropriate connection obtains between the rightness of an action and its agent’s motivation such that the agent is creditworthy for doing the right thing. (Singh 2020: 158)

¹⁶⁹ For criticism of this way of viewing accidentality, see Isserow (2021) and Markovits (2010).

Briefly, the worry is that if we were to accept this explanation of accidentality, it would seem to follow that no one acts in a morally worthy manner.

Finally, some allude to an epistemic condition. For instance, Zoë Johnson King defends the following account of accidentality:

CLAIM: for all types of acts *A*, someone accidentally *As* if she has no idea that she is performing an act of type *A* when she does so. (Johnson King 2020: 196)

It is not my part of my brief here to try to settle the dispute over how we should explicate the non-accidentality constraint. However, I do wish to note that the view that mirrors the wrong-making view** will have problems with Johnson King's explication of accidentality. Consider the case of Huck again (§3.3). He believed that it was wrong to help Jim escape. The view that mirrors the wrong-making view** implies that Huck is praiseworthy for helping Jim escape, since he helped Jim willingly while (presumably) being aware of the facts that constitute normative reasons which, taken together, make it right for him to help Jim flee. Now, consider Johnson King's explication again:

CLAIM: for all types of acts *A*, someone accidentally *As* if she has no idea that she is performing an act of type *A* when she does so. (Johnson King 2020: 196)

Although Huck was aware of the right-making features, and although he was motivated by those features, he had no idea that they were features that made it right for him to help Jim escape, and no idea he was acting rightly. Therefore, granting CLAIM, his action was just an accident, which means he was not acting in a praiseworthy way in helping Jim (granting that one cannot be praiseworthy for accidents). According to Johnson King, this verdict about Huck

shows that being motivated to do the right thing by the feature that makes it right is insufficient for [positive] moral worth. (Johnson King 2020: 197)

Note that the fact that Huck is not praiseworthy does not entail that it is not fitting to admire him, or that he lacks good or "praiseworthy" character traits. As Johnson King writes:

Huckleberry Finn has a praiseworthy character trait: he cares about Jim. And this led him to perform an act of a good type: it is morally right. But it does not follow that Huck is praiseworthy for performing an act of this type. And, in fact, he is *not* praiseworthy for performing an act of this type. Huck accidentally does the right thing, and we are not praiseworthy for that which we do accidentally. So, Huck

lacks the particular kind of praiseworthiness that is the mark of an act with moral worth. (Johnson King 2020: 200)

Again, I will not take a stance here on whether Johnson King's explication of the non-accidentality constraint is feasible. My aim is merely to show that on some current explications of the constraint, the view that mirrors the wrong-making view** view entails that accidents can be praiseworthy. If a view mirroring the wrong-making view** is to be at least *prima facie* plausible then this point would have to be addressed together with the intuition that people are praiseworthy for having performed (high stakes) supererogatory actions and omissions, and not merely right actions and omissions.

8.3 On Standing to Praise

In Chapters 4 and 5, I relied on the distinction between private and overt blame. As hinted in §8.1, it seems feasible to make a similar distinction with regard to praise. In other words, assuming that praise, similarly to blame, is a sentiment, we can claim that we praise someone privately when we merely harbour the sentiment of praise. This kind of praise does not involve interacting with its target. An agent can have the sentiment of praise and the desire to applaud the praiseworthy agent without acting overtly on that desire. When we (sincerely) praise someone overtly, in contrast, we have the sentiment of praise and act on its manifestations. For example, we act on the desire to applaud the praiseworthy agent by applauding.

Further, given the similarities between blame and praise, the features distinguishing private and overt blame I have highlighted in earlier chapters are relevant to the matter of private and overt praise as well. For one thing, private and overt praise aggregate differently. And we also need to account for their degrees differently.

Turning to standing to praise, while the idea of standing to blame is connected with well-known verbal responses such as "Who are you to blame me?", it is unclear what familiar responses of this kind, if any, are associated with the idea of standing to praise someone. This observation casts doubt on the claim that we have the intuition that a certain standing is a prerequisite of praising someone.

That said, I believe we can construct cases showing that we do in fact have this intuition. Consider the following case provided by Telech (2022):

[S]uppose Cleo praises Alicia for helping Bader though Cleo herself lacks commitment to the value underlying Alicia's praiseworthiness, as evinced by Cleo's regularly failing (without sufficient reason) to take on *pro bono* cases aiding socioeconomically underserves persons. Though merited and satisfactory of the epistemic condition, we might nevertheless think that, owing to Cleo's lack of commitment to value underlying Alicia's praiseworthy act, Cleo lacks the standing to praise Alicia. There is something *hollow* about Cleo's praise, given her lack of commitment to the relevant values, such that Alicia would understandably be put off by praise from Cleo. We need not assume Cleo's praise is *insincere*, only that it is improperly grounded in her evaluative dispositions (where these might be understood to consist in attentive, affective, and deliberative dispositions concerning the value in question). On a communicative view of praise, Cleo is not in a position to properly host (Telech, 2021, 172). Cleo is in one respect comparable to the fair-weather fan, who partakes in celebration of a team's victories, but who supports that team only when it is succeeding. Cleo, furthermore – at least according to the communicative view – *directs* Alicia (albeit, invitationally) to participate in jointly valuing Alicia's action. Given the superficiality of Cleo's relation to the value of providing legal aid to the underserved, Alicia may be in a position to respond by disregarding or discounting this directive. (Telech 2022: 11)

Telech is relying here on the commitment view. This states, as will be recalled from Chapter 4 (§4.2), that hypocritical blame indicates a lack of genuine commitment to *N* (a moral norm) on the blamer's part. It is this lack of commitment to *N* that undermines the blamer's standing to blame another for violating *N*. Granting such a view, it is easy to imagine cases in which an agent is insufficiently committed to a certain norm in a way that explains why she lacks standing to praise someone.¹⁷⁰

Let us grant that we need to have a certain standing in order to praise someone. Would this suggestion, comparably to the case of blame, only apply to overt praise? Or does it clearly apply to private praise as well? I think that we have reasons to be sceptical about the claim that there are norms of standing to praise

¹⁷⁰ For other cases showing the need of a standing to praise, see, in particular, Lippert-Rasmussen (2022) and Jeppsson and Brandenburg (2022).

privately. The reasons are similar to those we had for being sceptical about there being standing to blame privately,

If we believe the nature of standing is a normative power of some sort, standing does not apply to private praise. We do not exercise a normative power by having the sentiment of praise. However, if we think standing denotes a privilege, or if we adopt a mixed view of some sort with privilege as one element, standing will apply to private praise.

What about the specific defeaters of standing? I have the feeling that the avoidance of meddling condition (which treats facts showing that one is an outsider *vis-à-vis* the underlying matter as inimical to standing) does not apply to praise at all. Consider a case in which my neighbour's child has performed a praiseworthy action – say, saved a stranger from danger at great cost to herself. I do not have a close relationship with my neighbour or her child. Nor am I in any way affected by the child's action, and so forth. Here, my sense is that I have standing to praise the child both overtly and privately. I would not wrong the child if I praised her.

Moving on to the absence of hypocrisy condition, I concede that there is something amiss about agents who, for example, praise themselves privately for having φ -ed but never praise others for φ -ing. Intuitively, such agents lack standing to praise themselves privately. But, as with private blame, I find it hard to see that such agents are doing anything *pro tanto* wrong by privately praising only themselves (granting that they are praising themselves for the “right” kinds of reason and cannot directly stop praising themselves privately because it is hypocritical to do so, etc.). Therefore, it is not clear to me that hypocrisy entails that one loses one's standing, or privilege, to praise oneself privately.

Finally, familiar worries about guidance apply to private praise as well. Since the sentiment of praise is presumably a response to the “right” kinds of reason, and since reasons appealing to standing are presumably not of the right kind, we cannot be guided *directly* by specific conditions of standing. One cannot directly cease to feel gratitude to someone, or oneself, for the reason that it would be hypocritical to have that feeling. So, it would appear that we can only cease to privately praise another *indirectly* for this kind of reason – e.g., by going to therapy, through hypnosis, taking a pill, or what have you. And it is clearly a disadvantage of the theory that, if it were correct, meeting the conditions of standing would require one to do very demanding things. What makes these

methods too demanding is the fact that they oblige the agent to become irrational, or less reason-responsive, through some indirect atypical means.

Interestingly, and as with overt blame, none of these worries apply to overt praise. We can exercise a power by praising someone overtly, and we easily can readily refrain from praising someone overtly for the reason that it would be (say) hypocritical.

In sum, there do seem to be asymmetries between norms of standing to praise and norms of standing to blame. However, these differences do not prove that there is no such thing as having, or lacking, the standing to praise someone. Actually, we seem to have a reason to believe that there is such a thing (recall the case provided by Telech above). That said, I believe that, for reasons similar to those showing why we should be sceptical about the idea that there are norms of standing to blame privately, we should also be sceptical about the claim that there are norms of standing to praise privately.

8.4 On Proportional Praise

It is generally agreed that disproportionate blame (at least, disproportionate *overt* blame, see §5.2.1) is morally problematic in some way, and that we should incorporate a condition of proportionality in the ethics of blame. However, I believe that the same cannot be said about praise. There seems to be nothing unjust, or morally problematic, about praising someone to a degree that is, for example, excessive given her degree of praiseworthiness – especially when she is not harmed by such praise and when the praise is not unequally distributed.¹⁷¹

Granting that there is something morally problematic about disproportionate praise, and that we should incorporate a condition about proportionality in the ethics of praise, I now turn, briefly, to investigate whether views that mirror the proportionality condition for private and overt blame are plausible.

It is natural to hold that private praise is proportionate if, and only if, its degree matches the degree of private praise that it is fitting to direct upon the praiseworthy agent (i.e., the agent's degree of praiseworthiness), just as it is natural

¹⁷¹ For a good discussion on praise and fairness, see King (2014).

to think that the degree of other attitudes should correspond to the degrees of the relevant value.¹⁷² It is also natural to suppose that we make sense of degrees of private praise, as we make sense of other attitudes' degrees, in terms of their intensity. The more intense my private praise is, the more I praise the relevant certain agent (and *vice versa*). Thus, the proportionality condition for private blame seems readily translatable to private praise. We can formulate it as follows:

Proportional private praise: an instance of privately praising praiseworthy agent *B* for having ϕ -ed with intensity *i* is proportionate if, and only if, *i* matches the intensity *i* of private praise that it is fitting to direct upon *B* for having ϕ -ed.

Disproportional private praise: an instance of privately praising a praiseworthy agent *B* for having ϕ -ed with intensity *i* is disproportionate if, and only if, *i* exceeds, or is less than, the intensity *i* of private praise that it is fitting to direct upon *B* for having ϕ -ed.

As with disproportionate private blame (§5.2.1), there seems to be nothing morally problematic about experiencing (say) a too intense feeling of praise, just as it does not seem to be anything morally problematic about experiencing a too intense feeling of love, admiration or private blame.

In Chapter 5, §5.2.2, I presented two ways of making sense of proportional overt blame:

*Disproportionate overt blame conveyed**: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed to a certain degree_{conveyed} *n* is disproportionate_{conveyed} if, and only if, *n* exceeds, or is less than, the intensity *i* of private blame that it is fitting to direct upon *B* for having ϕ -ed.

*Disproportionate overt blame harm**: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed to a certain degree_{harm} *n* is disproportionate_{harm} if, and only if, *n* exceeds the goodness brought about by the overt blame in terms of *B* taking responsibility for having ϕ -ed.

*Proportionate overt blame conveyed**: an act of overtly blaming a blameworthy agent *B* for having ϕ -ed to a certain degree_{conveyed} *n* is

¹⁷² How we make sense of degrees of praiseworthiness, and what makes it fitting to praise someone privately to degree *n*, is a harder question, and one I will not attempt to answer here.

proportionate_{conveyed} if, and only if, n matches the intensity of private blame that it is fitting to direct upon B for having ϕ -ed.

*Proportionate overt blame harm**: an act of overtly blaming a blameworthy agent B for having ϕ -ed to a certain degree_{harm} n is proportionate_{harm} if, and only if, n does not exceed the goodness brought about by the overt blame in terms of B taking responsibility for having ϕ -ed.

Is a view mirroring the above principles reasonable when it comes to overt praise? To answer this question, we need to know more about how we should make sense of degrees of overt praise and the legitimate goals of overt praise.

Starting with degrees, as with overt blame, we could reasonably look at the intensity of the private praise the overt praise conveys – the more intense the private praise, the higher the degree of overt praise. Another suggestion, and one that mirrors the harm-way of making sense of overt blame's degree, is that we should look at the *benefits* the overt praise confers. The greater the benefits of one's overt praise for the praiseworthy agent, the higher the degree of one's praise (and *vice versa*). If we assume a hedonistic account, we can claim, more precisely, that the more pleasure one's overt praise gives to the praiseworthy agent, the more one praises her overtly (and *vice versa*).

In addition, it is relevant here to distinguish between narrow and wide proportionality (i.e., between benefits for the praiseworthy agent and benefits also for those with ties to her). Here I have in mind the narrow sense.

Similar aggregation problems as those concerning blame can arise in connection with praise as well. Suppose several people praise me for helping the lady up the stairs. Each act of overt praise benefits me to the same low degree. However, all of the praise, taken together, benefits me a lot.

Moving on to the good of praising, the relevant goods of overt praise include, but are not limited to, encouraging the praiseworthy agent to continue to do what she is doing at the time, and celebrating her on the basis of the act she has performed.

With these rough-and-ready explanations to hand, it is easier to assess whether a view of praise that mirrors the proportionality condition for overt blame is attractive. Specifically, are the following conditions plausible?

Disproportionate overt praise benefit: an act of overtly praising a praiseworthy agent B for having ϕ -ed to a certain degree_{benefit} n is disproportionate_{benefit} if, and only if, n exceeds the goodness brought about by the overt praise in terms of (say) celebrating B for having ϕ -ed.

Disproportionate overt praise conveyed: an act of overtly praising a praiseworthy agent B for having ϕ -ed to a certain degree_{conveyed} n is disproportionate_{conveyed} if, and only if, n exceeds, or is less than, the intensity i of private praise that it is fitting to direct upon B for having ϕ -ed.

Proportionate overt praise benefit: an act of overtly praising a praiseworthy agent B for having ϕ -ed to a certain degree_{benefit} n is proportionate_{benefit} if, and only if, n does not exceed the goodness brought about by the overt praise in terms of (say) celebrating B for having ϕ -ed.

Proportionate overt praise conveyed: an act of overtly praising a praiseworthy agent B for having ϕ -ed to a certain degree_{conveyed} n is proportionate_{conveyed} if, and only if, n matches the intensity i of private praise that it is fitting to direct upon B for having ϕ -ed.

Although these proportionality principles for overt praise are a bit odd, they are not obviously false. For example, there is something intuitive about the claim that it is disproportionate to praise a praiseworthy agent overtly in a way that involves benefitting her to a degree that exceeds the goodness of achieving the goal of (say) celebrating her on the basis of an action or omission she has performed. However, to evaluate these principles further and to answer the question whether disproportionate_{benefit/conveyed} overt praise is morally problematic, is a task for future research.

8.5 On Fitting Praise

In Chapter 6, I argued that the FA analysis of blameworthiness should be restricted so that it appeals only to fitting private blame. Does a similar argument apply to the FA analysis of praiseworthiness? Should we even try to formulate an FA analysis of praiseworthiness?

I believe that we easily can formulate such an account, as praiseworthiness, like blameworthiness, is a value property. For example, like other value properties, praiseworthiness admits of degrees.¹⁷³

If the FA analysis of praiseworthiness analyses praiseworthiness in terms of fitting private praise *and* overt praise, we would run into problems with standing and (perhaps) proportionality similar to those we encountered in connection with blame, since, as I have shown in §8.3-§8.4, there seem to be norms of standing and (perhaps) proportionality for praise that mirror the norms of standing and proportionality for blame.

Finally, given that we can make parallel distinctions regarding private and overt praise/blame, we can argue, for the same reasons as those provided in Chapter 6, that the FA analysis of praiseworthiness should be restricted so that it appeals only to fitting private praise. With that restriction, we will avoid the challenges from standing and (perhaps) proportionality.

8.6 Summary

Praise and blame share some important features. It is quite reasonable to think that they are more or less two sides of the same coin. However, there are some important disanalogies between them. It seems that the things we demand from agents if they are to be praiseworthy for the actions or omissions they perform and the things we demand from them if they are to be blameworthy for the actions, or omissions, they perform are different. And some familiar responses, such as “Who are you to blame me?”, are connected with the idea of standing to blame but not with the idea of the standing to praise. The dissimilarities suggest we cannot straightforwardly generalise what I have said about blame to praise. Praise requires its own treatment.

¹⁷³ For an FA analysis of praiseworthiness, see King (2012).

9 References

- Aciman, A. (2017). *Call Me By Your Name*. Atlantic Books.
- Alvarez, M. & Littlejohn, C. (2017). When Ignorance is No Excuse. In Robichaud, P. & Wieland, J. W. (Eds.). *Responsibility – The Epistemic Condition*. Oxford University Press, 64-81.
- Andersson, H., & Werkmäster, J. G. (2020). How Valuable Is It? *Journal of Value Inquiry* Vol 3: 1-18.
- Andersson, H: & Werkmäster, J. (2022). Normative Resilience. *Utilitas* Vol. 34 (2): 195-208.
- Arneson, R. (2003). The Smart theory of moral responsibility and desert. In Olsaretti, S. (Ed.). *Desert and Justice*. Clarendon Press.
- Arpaly, N. (2002). Moral Worth. *Journal of Philosophy* Vol. 99 (5): 223-245.
- Arpaly, N., & Schroeder, T. (2013). *In Praise of Desire*: Oxford University Press.
- Ayars, A. A. (2021). Blaming for Unreasonableness: Accountability Without Ill Will. *Journal of Ethics and Social Philosophy* Vol. 19 (1): 56-79.
- Basu, R. & Schroeder, M. (2019). Doxastic Wronging. In Kim, B. & McGrath, M. (Eds.). *Pragmatic Encroachment in Epistemology*. Routledge, 181-205.
- Ben-Ze'ev, A. (2000). *The Subtlety of Emotions*. Vol. 63. Bradford.
- Berker, S. (2022). The Deontic, the Evaluative, and the Fitting. In Howard, C. & Cosker-Rowland, R. (Eds.). *Fittingness: Essays in the Philosophy of Normativity*. Oxford University Press, 23-57.
- Billingham, P., & Parr, T. (2020a). Enforcing social norms: The morality of public shaming. *European Journal of Philosophy* Vol. 28 (4): 997-1016.
- Billingham, P., & Parr, T. (2020b). Online Public Shaming: Virtues and Vices. *Journal of Social Philosophy* Vol. 51(3): 371-390.
- Björnsson, G. (2017). Explaining Away Epistemic Scepticism about Culpability. In Shoemaker, D. (Ed.). *Oxford Studies in Agency and Responsibility Volume 4*. Oxford University Press, 141-164.

- Brekke, Carlsson, A. (2017). Blameworthiness as Deserved Guilt. *Journal of Ethics* Vol. 21 (1): 89-115.
- Brekke Carlsson, A. (2019). Shame and Attributability. In Shoemaker, D. (Ed.). *Oxford Studies in Agency and Responsibility*, vol. 6, 112-139.
- Bruno, D. (2022). Being Fully Excused For Wrongdoing. *Pacific Philosophical Quarterly* Vol. 104 (2):1-24.
- Bykvist, K. (2009). No Good Fit: Why the Fitting Attitude Analysis of Value Fails. *Mind* Vol. 118 (469): 1-30.
- Bykvist, K. (2018). Agent-Relative and Agent-Neutral Reasons. In Star, D. (Ed.). *The Oxford Handbook of Reasons and Normativity*. Oxford University Press, 821-838.
- Clarke, R. (2017). Blameworthiness and Unwitting Omissions. In Nelkin, D. K. & Rickless, S. C.. (Eds.). *The Ethics and Law of Omissions*. Oxford University Press, 63-83.
- Clarke, R. (forthcomin). Accounting for Failure. In Cyr, T., Law, A. & Tognazzini, N. (Eds.). *Freedom, Responsibility, and Value: Essays in Honour of John Martin Fischer*. Routledge, 153-170.
- Coates, D. J. (2019). Being More Blameworthy. *American Philosophical Quarterly* Vol. 56 (3): 233-246.
- Coates, D. J. (2020). The Ethics of Blame: A Primer. In Ernst, G. & Schmidt, S. (Eds.). *The Ethics of Belief and Beyond. Understanding Mental Normativity*. Routledge, 192-214.
- Cohen, G. A. (2006). Casting the First Stone: Who Can, and Who Can't, Condemn the Terrorists? *Royal Institute of Philosophy Supplement* Vol. 58: 113-136.
- Cosker-Rowland, R. (2019). *The Normative and the Evaluative: The Buck-Passing Account of Value*. Oxford University Press.
- Crisp, R., & Cowton, C. J. (1994). Hypocrisy and Moral Seriousness. *American Philosophical Quarterly* Vol. 31 (4): 343-349.
- Crisp, R. (2000). Value ... And what follows by Joel Kupperman new York: Oxford university press, £25.00. *Philosophy* Vol. 75 (3): 452-462.
- Crisp, R. (2005). Value, reasons and the structure of justification: how to avoid passing the buck. *Analysis* Vol. 65 (1): 80-85.
- D'Arms, J., & Jacobson, D. (2000a). The Moralistic Fallacy: On the 'Appropriateness' of Emotions. *Philosophical and Phenomenological Research* Vol. 61 (1): 65-90.
- D'Arms, J., & Jacobson, D. (2000b). Sentiment and Value. *Ethics* Vol. 110: 722-748.
- D'Arms, J. (2003). VIII. The significance of recalcitrant emotion (or, anti-quasijudgmentalism). *Royal Institute of Philosophy Supplement* Vol 52: 127-145.

- Danielsson, S., & Olson, J. (2007). Brentano and the Buck-Passers. *Mind* Vol. 116 (463): 511-522.
- Darwall, S. L. (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Harvard University Press.
- Deonna, J., & Teroni, F. (2009). Taking Affective Explanations to Heart. *Social Science Information* Vol. 48 (3): 359-377.
- Deonna, J., & Teroni, F. (2011). *The Emotions: A Philosophical Introduction*. Routledge.
- Edlich, A. (2022). What About the Victim? Neglected Dimensions of the Standing to Blame. *The Journal of Ethics* Vol. 26 (2): 209-228.
- Edwards, J. (2019). Standing to Hold Responsible. *Journal of Moral Philosophy* Vol. 16 (4): 437-462.
- Enoch, D., & Marmor, A. (2007). The case against moral luck. *Law and Philosophy* Vol. 26 (4): 405-436.
- Enoch, David & Spectre, Levi (forthcoming). There is no such thing as doxastic wrongdoing. *Philosophical Perspectives*.
- Fricker, M. (2016). What's the Point of Blame? A Paradigm Based Explanation. *Noûs* Vol. 50 (1): 165-183.
- Fritz, K. G., & Miller, D. (2018). Hypocrisy and the Standing to Blame. *Pacific Philosophical Quarterly* Vol. 99 (1): 118-139.
- Fritz, K. G., & Miller, D. J. (2019). The Unique Badness of Hypocritical Blame. *Ergo* Vol. 6 (19): 545-569.
- Fritz, K. G., & Miller, D. J. (2022). A Standing Asymmetry between Blame and Forgiveness. *Ethics* Vol. 132 (4): 759-786.
- Frowe, H. (2011). *The Ethics of War and Peace: An Introduction*. Routledge.
- Garcia, A. (2018). *Between Values and the World: Studies in second-order value theory*. Media-Tryck.
- Gertken, J., & Kiesewetter, B. (2017). The right and the wrong kind of reasons. *Philosophy Compass* Vol. 12 (5): 1-14.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. (Vol. 41). Harvard University Press.
- Goldie, P. (2011a). Grief: A narrative account. *Ratio* Vol. 24 (2): 119-137.
- Goldie, P. (2011b) Intellectual Emotions and Religious Emotions. *Faith and Philosophy: Journal of the Society of Christian Philosophers* Vol. 28 (1): 93-101.
- Graham, P. A. (2014). A Sketch of a Theory of Moral Blameworthiness. *Philosophy and Phenomenological Research* Vol. 88 (2): 388-409.

- Haji, I. (1997). An Epistemic Dimension of Blameworthiness. *Philosophy and Phenomenological Research* Vol. 57(3): 523-544.
- Hartman, Robert J. (2019). Moral Luck and The Unfairness of Morality. *Philosophical Studies* Vol. 176 (12): 3179-3197.
- Heil, J. (2003). *From an Ontological Point of View*. Oxford University Press.
- Helm, B. W. (2009). Love, identification, and the emotions. *American Philosophical Quarterly* Vol. 46 (1): 39-59.
- Herstein, Ori J. (2020). Justifying Standing to Give Reasons: Hypocrisy, Minding Your Own Business, and Knowing One's Place. *Philosophers' Imprint* Vol. 20 (7): 1-18.
- Heuer, U. (2010). Beyond Wrong Reasons: The Buck-Passing Account of Value. In Brady, M. (Ed.). *New Waves in Metaethics*. Palgrave-Macmillan, 166-184.
- Hieronymi, P. (2004). The force and fairness of blame. *Philosophical Perspectives* Vol. 18 (1): 115–148.
- Hieronymi, P. (2005). The Wrong Kind of Reason. *Journal of Philosophy* Vol. 102 (9): 437 - 457.
- Hieronymi, P. (2008). Review: Sher's Defense of Blame. *Philosophical Studies* Vol. 137 (1): 19-30.
- Howard, N. R. (forthcoming). The Goals of Moral Worth. *Oxford Studies in Metaethics*.
- Hume, D. (1738). *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning Into Moral Subjects*. Oxford University Press.
- Inwagen, P. V. (1975). The Incompatibility of Free Will and Determinism. *Philosophical Studies* Vol. 27 (3): 185 - 199.
- Isserow, J. (2021). Doubts about Duty as a Secondary Motive. *Philosophy and Phenomenological Research* Vol. 105 (2): 276-298.
- Jackson, F. (1991). Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics* Vol. 101 (3): 461-482.
- James, W. (1884). What is an Emotion? *Mind* Vol. 9 (34): 188-205.
- Jeppsson, Sofia & Brandenburg, Daphne. (2022). Patronizing Praise. *The Journal of Ethics*. Vol. 26 (4): 663-682.
- Johansson Werkmäster, M. (2022). Blame as a sentiment. *International Journal of Philosophical Studies* Vol. 3: 1-15.
- Johnson King, Z. (2020). Accidentally Doing the Right Thing. *Philosophy and Phenomenological Research* Vol. 100 (1): 186-206.
- Kiesewetter, B. (2017). *The Normativity of Rationality*. Oxford University Press.

- King, M. (2012). Moral Responsibility and Merit. *Journal of Ethics and Social Philosophy* Vol. 6. (2): 1-17.
- King, M. (2014). Two faces of desert. *Philosophical Studies* Vol. 169 (3): 401-424.
- King, M. (2019a). Skepticism About the Standing to Blame. In Shoemaker, D. (Ed.). *Oxford Studies in Agency & Responsibility*. Vol. 6. Oxford University Press, 265-288.
- King, M. (2019b). Attending to blame. *Philosophical Studies* Vol. 177(5): 1423-1439.
- Kubala, R. (2017). Valuing and believing valuable. *Analysis* Vol. 77 (1): 59-65.
- Lee, Y. (2012). Why Proportionality Matters. *University of Pennsylvania Law Review* Vol. 160(6): 1835-1852.
- Lemos, N. (2011). Intrinsic Value and the Partiality Problem. *Philosophy and Phenomenological Research* Vol. 82 (3): 697-716.
- Lippert-Rasmussen, K. (2020). Why the Moral Equality Account of the Hypocrite's lack of Standing to Blame Fails. *Analysis* Vol. 80 (4): 666-674.
- Lippert-Rasmussen, K. (2022). Praising Without Standing. *The Journal of Ethics*. Vol. 26 (2): 1-18.
- Lord, E. (2017). What You're Rationally Required to Do and What You Ought to Do (Are the Same Thing!). *Mind* Vol. 126: 1109-1154.
- Macnamara, C. (2015). Blame, Communication, and Moral Responsible Agency. In Clarke, R, McKenna, M. & Smith, A. (Eds.). *The Nature of Moral Responsibility: New Essays*. Oxford University Press, 211-236.
- Manley, D., & Wasserman, R. (2008). On linking dispositions and conditionals. *Mind* Vol. 117 (465): 59-84.
- Markovits, J. (2010). Acting for the right reasons. *Philosophical Review* Vol. 119 (2): 201-242.
- Markovits, J. (2012). Saints, heroes, sages, and villains. *Philosophical Studies* Vol. 158 (2): 289-311.
- Marušić, B. (2018). Do Reasons Expire? An Essay on Grief. *Philosophers' Imprint* Vol. 18 (25): 1-21.
- Marušić, B. (2022). *On the Temporality of Emotions: An Essay on Grief, Anger, and Love*. Oxford University Press.
- Mason, E. (2019). *Ways to Be Blameworthy: Rightness, Wrongness, and Responsibility*. Oxford University Press.
- McHugh, C., & Way, J. (2016). Fittingness First. *Ethics* Vol. 126 (3): 575-606.
- McHugh, C. (2017). Attitudinal control. *Synthese* Vol. 194 (8): 2745-2762.

- McHugh, C. & Way, J. (2022). *Getting Things Right: Fittingness, Value and Reasons*. Oxford University Press.
- McKenna, M. (2012). *Conversation & Responsibility*. Oxford University Press.
- McMahan, J. (2014). Proportionality and Just Cause. *Journal of Moral Philosophy* Vol. 11 (4): 428-453.
- Menges, L. (2017). The emotion account of blame. *Philosophical Studies* Vol. 174 (1): 257-273.
- Menges, L. (forthcoming). Blaming. In Kiener, M. (Ed.). *The Routledge Handbook of Responsibility*.
- Milam, P. E. (2021). Get Smart: Outcomes, Influence, and Responsibility. *The Monist* Vol. 104 (4): 443-457.
- Miller, D. J. (forthcoming). The Epistemic Condition. In Kiener, M. (Ed.). *The Routledge Handbook of Moral Responsibility*. Routledge.
- Molnar, G. (2003). Mumford, S. (Ed.). *Powers: A Study in Metaphysics*. Oxford University Press.
- Murray, S. (2017). Responsibility and vigilance. *Philosophical Studies* Vol. 174: 507-527.
- Naar, H. (2013). A Dispositional Theory of Love. *Pacific Philosophical Quarterly* Vol. 94 (3): 342-357.
- Naar, H. (2018). Sentiments. In Naar, H. & Teroni, F. (Eds.). *The Ontology of Emotions*. Cambridge University Press.
- Nagel, T. (1979). *Moral Questions*. Cambridge University Press.
- Nelkin, D. K. (2016). Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness. *Noûs* Vol. 50 (2):356-378.
- Nussbaum, M. C. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press.
- Nussbaum, M. C. (2016). *Anger and Forgiveness: Resentment, Generosity, and Justice*. Oxford University Press.
- Olson, J. (2009). Fitting Attitude Analyses of Value and the Partiality Challenge. *Ethical Theory and Moral Practice* Vol. 12(4): 365-378.
- Olson, J. (2018). The Metaphysics of Reasons. In Star, D. (Ed.). *The Oxford Handbook of Reasons and Normativity*. Oxford University Press, 255-274.
- Orsi, F. (2013). Fitting Attitudes and Solitary Goods. *Mind* Vol. 122 (487): 687-698.
- Orsi, F. (2015). *Value Theory*. Bloomsbury Academic.
- Parfit, D. (2011). *On What Matters: Two-Volume Set*. Oxford University Press.

- Pereboom, D. (2009). Free Will, Love and Anger. *Ideas Y Valores* Vol. 58 (141): 169-189.
- Pickard, H. (2013). Irrational blame. *Analysis* Vol. 73(4): 613-626.
- Portmore, D. W. (2022). A Comprehensive Account of Blame: Self-Blame, Non-Moral Blame, and Blame for the Non-Voluntary. In Carlsson, A. B. (Ed.). *Self-Blame and Moral Responsibility*. Cambridge University Press, 48-76.
- Prinz, J. (2004). *Gut Reactions: A Perceptual Theory of the Emotions*. Oxford University Press.
- Rabinowicz, W. (2013). Value, Fitting-Attitude Account of. In LaFollette, H. (Ed.). *The International Encyclopedia of Ethics*. Wiley-Blackwell, 1-12.
- Rabinowicz, W., & Rønnow-Rasmussen, T. (2004). The strike of the demon: On fitting pro-attitudes and value. *Ethics* Vol. 114(3): 391-423.
- Rabinowicz, W., & Rønnow-Rasmussen, T. (2006). Buck-passing and the right kind of reasons. *Philosophical Quarterly* Vol. 56(222): 114-120.
- Reisner, A. E. (2009). Abandoning the Buck Passing Analys of Final Value. *Ethical Theory and Moral Practice* Vol. 12: 379-395.
- Reisner, A. E. (2015). Fittingness, Value and trans-World Attitudes. *Philosophical Quarterly* Vol. 65 (260): 1-22.
- Rivera-lópez, E. (2006). Can There Be Full Excuses for Morally Wrong Actions? *Philosophy and Phenomenological Research* Vol. 73(1): 124-142.
- Roadevin, C. (2018). Hypocritical Blame, Fairness, and Standing. *Metaphilosophy* Vol. 49 (1-2): 137-152.
- Rønnow-Rasmussen, T. (2011). *Personal Value*. Oxford University Press.
- Rosen, G. (2004). Skepticism about moral responsibility. *Philosophical Perspectives* Vol. 18(1): 295-313.
- Rosen, G. (2008). Kleinbart the Oblivious and Other Tales of Ignorance and Responsibility. *Journal of Philosophy* Vol. 105 (10): 591-610.
- Rosen, G. (2015). The Alethic Conception of Moral Responsibility. In Clarke, R., McKenna, M. & Smith, A. (Eds.). *The Nature of Moral Responsibility: New Essays*. Oxford University Press, 65-88.
- Rudy-Hiller, R. (2017). A Capacitarian Account of Culpable Ignorance. *Pacific Philosophical Quarterly* Vol. 98 (1): 398-426.
- Scanlon, T. (1998). *What We Owe to Each Other*. (Vol. 66). Harvard University Press.
- Scanlon, T. (2008). *Moral Dimensions: Permissibility, Meaning, Blame*. Harvard University Press.

- Scanlon, T. (2013). Interpreting Blame. In Coates, J. & Tognazzini, N. A. (Eds). *Blame: Its Nature and Norms*. Oxford University Press, 84-99.
- Scheffler, S. (2010). *Equality and Tradition: Questions of Value in Moral and Political Theory*. Oxford University Press.
- Schroeder, M. (2007). Teleology, agent-relative value, and 'good'. *Ethics* Vol. 117 (2): 265-295.
- Schroeder, M. (2010). Value and the right kind of reason. *Oxford Studies in Metaethics Volume 5*. Oxford University Press, 25-55.
- Schroeder, M. (2012). The Ubiquity of State-Given Reasons. *Ethics* Vol. 122 (3): 457-488.
- Sharadin, N. (2016). Reasons Wrong and Right. *Pacific Philosophical Quarterly* Vol. 97(3): 371-399.
- Sher, G. (2005). *In Praise of Blame*. Oxford University Press.
- Sher, G. (2021). *A Wild West of the Mind*. Oxford University Press.
- Shoemaker, D. (2015). *Responsibility From the Margins*. Oxford University Press.
- Shoemaker, D., & Vargas, M. (2019). Moral torch fishing: A signaling theory of blame. *Noûs* Vol. 55 (3): 581-602.
- Singh, K. (2020). Moral Worth, Credit, and Non-Accidentality. In Timmons, M. (Ed.), *Oxford Studies in Normative Ethics, Vol. 10*: Oxford University Press.
- Skorupski, J. (2010a). *The Domain of Reasons*. Oxford University Press.
- Skorupski, J. (2010b). Moral Obligation, Blame, and Self-Governance. *Social Philosophy and Policy* Vol. 27 (2): 158-180.
- Sliwa, P. (2016). Moral Worth and Moral Knowledge. *Philosophy and Phenomenological Research* Vol. 93 (2): 393-418.
- Sliwa, P. (2019). Reverse-engineering blame 1. *Philosophical Perspectives* Vol. 33 (1): 200-219.
- Sliwa, P. (2020). Excuse without Exculpation: The Case of Moral Ignorance. In Landau, R. S. (Ed.). *Oxford Studies in Metaethics Volume 15*. Oxford University Press, 72-95.
- Smith, A. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics* Vol. 115 (2): 236-271.
- Smith, A. (2007). On Being Responsible and Holding Responsible. *Journal of Ethics* Vol. 11(4): 465-484.
- Smith, A. (2013). Moral Blame and Moral Protest. In Coates, D. J. & Tognazzini, N. A.. (Eds.). *Blame: Its Nature and Norms*. Oxford University Press, 27-48.

- Smith, M. (1994). *The Moral Problem*. Blackwell.
- Snedegar, J. (forthcoming). Explaining Loss of Standing to Blame. *Journal of Moral Philosophy*.
- Solomon, R. C. (1976). *The Passions* (Vol. 87): University of Notre Dame Press.
- Stout, Nathan. (2020). On the significance of praise. *American Philosophical Quarterly* Vol. 57 (3): 215-226.
- Strawson, P. F. (1974 (1962)). Freedom and Resentment. In Strawson, P. F. (Ed.). *Freedom and Resentment and Other Essays*. Routledge, 1-28.
- Strawson, P. F. (1985). *Scepticism and Naturalism: Some Varieties*. (Vol. 36). Routledge.
- Tadros, V. (2018). Past Killings and Proportionality in War. *Philosophy and Public Affairs* Vol. 46 (1): 9-35.
- Talbert, M. (2012). Moral Competence, Moral Blame, and Protest. *The Journal of Ethics* Vol. 16 (1): 89-109.
- Talbert, M. (2017). Akrasia, Awareness, and Blameworthiness. In Robichaud, P. & Wieland, J. W. (Eds.). *Responsibility: The Epistemic Condition*. Oxford University Press, 47-63.
- Tappolet, C. (2016). *Emotions, Value, and Agency*. Oxford University Press.
- Telech, D., & Tierney, H. (2019). The Comparative Nonarbitrariness Norm of Blame. *Journal of Ethics and Social Philosophy* Vol. 16 (1): 25-43.
- Telech, D. (2022). Praise. *Philosophy Compass* Vol. 17 (10): 1-19.
- Thomson, J. J. (1990). *The Realm of Rights* (Vol. 42): Harvard University Press.
- Todd, P. (2017). A Unified Account of the Moral Standing to Blame. *Noûs* Vol 53: 347-374.
- Tognazzini, N. & Coates, J. (2021). Blame. *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2021/entries/blame/>>.
- Tolstoy, Leo. *Anna Karenina*. (Trans.). Volokhonsky, L. & Pevear, R. Penguin Classics.
- Tomlin, P. (2018). Subjective Proportionality. *Ethics* Vol. 129 (2): 254-283.
- Vendrell Ferran, Í. (2021). Hate: Toward a Four-Types Model. *Review of Philosophy and Psychology*: 1-19.
- Vetter, B. (2015). *Potentiality: From Dispositions to Modality*. Oxford University Press.
- Wallace, R. J. (1994). *Responsibility and the Moral Sentiments*. Harvard University Press.
- Wallace, R. J. (2010). Hypocrisy, Moral Address, and the Equal Standing of Persons. *Philosophy and Public Affairs* Vol. 38(4): 307-341.

- Wang, S. (2021). The Communication Argument and the Pluralist Challenge. *Canadian Journal of Philosophy* Vol. 51(5): 384-399.
- Watson, G. (2004). *Agency and Answerability: Selected Essays*. Oxford University Press.
- Werkmäster, J. (2019). *Reasons and Normativity*. Media-Tryck.
- Williams, B. (1981). *Moral Luck: Philosophical Papers 1973–1980*. Cambridge University Press.
- Woods, Jack. (2021). Ordinary Wrongdoing. In Timmons, M. (Ed.). *Oxford Studies in Normative Ethics*. Vol. 11. Oxford University Press, 155–175.
- Wolf, S. (2011). Blame, Italian Style. In Kumar, R. Wallace, R. J. & Freeman, S. (Eds.). *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. Oxford University Press, 332-346.
- Zimmerman, M. J. (1987). Luck and moral responsibility. *Ethics* Vol. 97 (2): 374-386.
- Zimmerman, M. J. (1988). *An Essay on Moral Responsibility*. Rowman & Littlefield.
- Zimmerman, M. J. (1997). Moral responsibility and ignorance. *Ethics* Vol. 107 (3): 410-426.
- Zimmerman, M. J. (2002). Taking luck seriously. *Journal of Philosophy* Vol. 99 (11): 553-576.
- Zimmerman, M. J. (2008). *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge University Press.
- Zimmerman, M. J. (2011a). *The Immorality of Punishment*. Broadview Press.
- Zimmerman, M. J. (2011b). Partiality and Intrinsic Value. *Mind* Vol. 120 (478): 447-483.
- Zimmerman, M. J. (2022). *Ignorance and Moral Responsibility*. Oxford University Press.

Aspects of Blame

This thesis discusses what blame is, what it is for an agent to be blameworthy for having performed a certain act or omission, what makes an agent blameworthy for having performed a certain act or omission, whether the idea of standing to blame generalises to blame understood as something privately held, and what it is for blame to be proportionate. It provides original answers to these questions that move the current discussion of blame forward.