

CONSORZIO  
di DOTTORATO  
in FILOSOFIA  
NORD-OVEST



UNIVERSITÀ DEGLI STUDI DI PAVIA – CONSORZIO FINO

TESI DI DOTTORATO DI RICERCA IN FILOSOFIA

THE NATURE OF BLAME

Supervisor:

Chiar.mo Professor Luca Fonnesu

Tesi di dottorato di:

Lorenzo Testa

CICLO XXXV



## ACKNOWLEDGMENT

Writing a doctoral dissertation is never an easy task. Writing a doctoral dissertation in the heat of a pandemic was even harder. Despite the harsh conditions, my doctoral studies have been populated by great people. I now turn to thank them and acknowledge the importance they have had for my journey.

First, I want to mention and thank my supervisor Luca Fonnesu, who followed my philosophical development since my bachelor's degree. At that time, I was working on free will and compatibilism. Then, during my master studies, I turned to the concept of moral responsibility. It was Luca Fonnesu who suggested I focused on blame, and I am indebted to him for this suggestion. His philosophical insights always helped me perfect my philosophical skills. I thank him for supporting my ideas about non-moral blame, which were sometimes seen as controversial during my years as a doctoral student.

Then, I wish to thank Sergio Filippo Magni for mentoring me during the first year of my doctoral study. I enjoyed our philosophical discussions, and I am deeply indebted to his suggestions about moral responsibility and blame. If I have started working on these topics, it is also thanks to him.

Thirdly, I want to thank Gunnar Björnsson, who welcomed me to Stockholm University as a visiting PhD student. His suggestions were precious, and his papers really helped me polishing my ideas about non-moral blame. Moreover, the months spent in Stockholm were an important life experience. I really wish to come back to Sweden in the near future.

Even if the pandemic made it difficult to build a strong relationship with my fellow PhD colleagues, I am indebted to every one of them for their suggestions during our Work in Progress Seminars. In no particular order, I wish to personally thank Matteo Boccacci, Giulia Cantamessi, Rita Ogochukwu Ezugwu, Matilde Liberti, Arianna Dini, Chiara Innorta, Luca Stroppa, Valerie Topf, Diego Alberto Biancolin, Carline Klijnman, Marco Miglino, Dario Cecchini, Jessica Lombard, Matteo Rategni, Steven Umbrello, and Laura Santi Amantini. A special thanks to Michele Bocchiola, Valeria Ottonelli, Corrado Fumagalli, Michel Croce, and Federico Zuolo for the organization of the Work in Progress seminars, which were among the best seminars I have had during my PhD years. Among professors, I am also indebted to Maria Silvia Vaccarezza, who commented on one of my first papers and helped me improve it.

Thanks to Tommaso Piazza and Ian Carter, with whom I had the opportunity to discuss my ideas about blame. Moreover, they are both keen supporters of the Pavia Debating Society, which I founded with my friends Sandra Innamorato, Lorenzo Botta Parandera, and Giulia Balossino in 2019. A special mention also goes to Alessandra Tanesini, with whom I had an interesting discussion not just about philosophy, but also about what it takes to be a professional philosopher. Last but not least, Carlo Penco for having invited me to participate in the birth of PhilVideos, which I hope will thrive in the future.

Turning to my friends, I shall quote someone who has endured a journey as though as earning a PhD: “I don't know half of you half as well as I should like; and I like less than half of you half as well as you deserve.” Thanks to Martina Stevanato, Giulia Milli, Marzia Marconi, Giorgio Matteoli, Simone Parizia, Paolo Rivoira, Luca Milanese, Pietro Arnolfo, Lorenzo Calzetta, Marco Villa, Francesco Burzio, Marco Uranio, Andrea Rinaudo, Alberto Lauro, Paolo Bianchi, Silvia Bianchi, Erika Breuza, and Marco & Alessandro for having lightened my heart during difficult times.

Among friends, a special mention goes to “Joker’s Friends” Paolo Gatti and Luca Passi. They are both important friends who read and commented on my philosophical works, always helping me. Their contribution, however, goes beyond philosophy. Underneath the surface of sarcasm and dark humor, Paolo and Luca really helped me when I was struggling. Paolo Gatti should also be mentioned alongside my dear friend Valeria Tavani. Probably, we disagree on most philosophical matters but may our friendship last forever.

My parents Silvano and Raffaella never stopped me from chasing my dreams. They have always supported my decision to pursue philosophical studies in my life, and for this reason I thank them. For the same reason, I also have to mention my uncle “Pio”.

Lastly, my most sincere gratitude goes to Alice, my partner during these last four years. Not only Alice has read (and re-read, and re-read...) this dissertation, but she has also always supported me in every possible way. Since June 2022, Annie came to our home and really enriched our life. Annie did not read this dissertation though, but not because she is a newborn: Annie is our cat.

## THE NATURE OF BLAME

### Contents

INTRODUCTION .....	2
CHAPTER ONE.....	7
UNDERSTANDING BLAME.....	7
§1 Blame, Causation, and Blameworthiness.....	8
§2 Framing the Debate on Blame.....	11
§3 Blame and Praise .....	19
§4 Moral Blame and Non-moral Blame .....	24
§5 What should a Theory of Blame Account for? .....	29
CHAPTER TWO.....	37
CONTENT-BASED THEORIES OF BLAME .....	37
§1 Cognitive Accounts of Blame .....	38
§2 Emotional Accounts of Blame .....	48
§3 Sher's Account of Blame as a Disposition Around a Belief-Desire Pair .....	55
§4 Scanlon's Relationship-based Account of Blame .....	62
CHAPTER THREE.....	72
FUNCTIONAL THEORIES OF BLAME.....	72
§1 Blame as Protest.....	74
§2 The Epistemic Function of Blame .....	86
§3 Communicative Blame: A Paradigm for our Blaming Practices? .....	94
CHAPTER FOUR.....	102
THE SIGNALING ACCOUNT OF BLAME.....	102
§1 The Hypothesis: Blame as a Signaling Device .....	103
§2 What does Signaling Mean? The Costly Signaling Theory .....	107
§3 Blame as Signaling.....	111
§4 The Data Points of Blame and Blame as Signaling .....	117
§5 Objections to Blame as Signaling.....	123
CHAPTER FIVE.....	129
NON-MORAL BLAME AS A FORM OF SIGNALING.....	129
§1 The Case against Non-Moral Blame: The Moralizing Trap.....	131
§2 Is Non-Moral Blame Impermissible?.....	136
§3 Non-moral Blame and the Signaling Theory of Blame .....	142
CONCLUSION .....	148

REFERENCES ..... 150



## INTRODUCTION

Blame is a pervasive phenomenon of our interpersonal life. We blame our friends when they show up late at an appointment, we blame dictators for their unjustified use of violence against innocents, we blame our teammate for a sloppy play, and we may even blame ourselves for our laziness or for other unappealing characteristics. All these instances of blame, in turn, can take place in different ways. We can directly and overtly address our latecomer friends, but probably we will never come face to face with the violent foreign dictator. We can blame absent people with those around us, and we may even blame the dead. Finally, our blame can be kept silent, as when we blame ourselves and others without expressing it.

As these few examples suggest, our blaming practices can take various forms. The variety of our blaming practices surely plays a role in any attempt to define the concept of blame. What is blame, and how should we describe our blaming practices? Is it possible to find a single element that ties together this striking variety? These questions are at the center of the so-called problem of the nature of blame. In this dissertation, I contribute to the debate over the nature of blame, attempting to find this unifying element that many kinds of theories of blame pursue.

The variety of our blaming practices is not the only difficulty linked to research on blame. Many philosophers discuss blame as part of the debate on free will. A common thought is that blame is only fitting (or, depending on the details of one's theory, deserved) on condition that an agent is free in the relevant sense. If determinism is true, and if the truth of determinism is not compatible with human agents being free, then no one would deserve to be blamed for her actions. Perhaps surprisingly, my dissertation does not directly deal with this kind of problem. As Justin Coates suggests, discussions of this kind treat blame in an



indirect sense. What is relevant in most debates over free will and moral responsibility, is not blame *per se*, but, rather, blameworthiness (Coates 2012, 197-198). But even when (or, I would say, if) we have understood everything about the conditions of blameworthiness, we will still lack an accurate description and explanation of the concept of blame. Is blame an emotion, or a cognitive mental state? Do blame and our blaming practices serve a function? As we can see, there are questions on the nature of blame that are not addressed by studies on free will and moral responsibility. The main aim of this dissertation is to analyze the concept of blame and to study our blaming practices, rather than contributing to the longstanding debate on free will.

Naturally, I am not alone in this project. In the past thirty years or so, many philosophers questioned blame as an important moral psychological phenomenon in its own right. Generically speaking, blame is a reaction to wrongdoing. Upon reflection, however, we should note that not *every* reaction to wrongdoing counts as blame. My spouse can act wrongly, and in response I may feel sad, disappointed, or maybe even amused (if I do not care about morality). But I would not say that these reactions count as blame. Thus, it seems that only *certain* reactions to wrongdoing count as blame. What are the conditions under which we can properly speak of blame? Does the blamer need to be in a specific mental state when she engages in her blaming activities, in order for those activities to count as blame? Or, perhaps, what is important to discriminate between blame and other reactions to wrongdoing is the function that blame serves?

My aim is to contribute to the contemporary debate on the nature of blame in two ways. First, I want to discuss and critically evaluate the most prominent attempts to answer the question: “What is blame?”. Second, I defend a theory on the nature of blame. According to this theory, blame is whatever mental state (cognitive, emotional, conative, or a mix of these) serves the function of *signaling* the blamer’s normative competence and normative

commitments. A version of the theory of blame as signaling has been recently defended by D. Shoemaker and M. Vargas in their 2021. To my knowledge, their work has not been yet amply discussed. However, I think they are on the right track when they argue that the most important function of blame is that of *signaling* something.

Finally, this dissertation also deals with a somewhat neglected topic in the philosophical discussion on blame. This topic is the status of non-moral blame. Does a theory of blame need to address cases of non-moral blame, too? Or, more radically, is non-moral blame to be considered as “proper” blame? In my opinion, a theory of blame that cannot account for cases of non-moral blame is at best incomplete. In the last chapter of this dissertation, I want to show that a theory of blame as signaling can also accommodate cases of non-moral blame.

This work is divided into five chapters. In the first chapter, I introduce the contemporary debate on blame, specifying the object of my research and laying the foundations of the rest of the chapters of this dissertation.

In the second and third chapter, I attempt a taxonomy of the most prominent contemporary works on the nature of blame. More specifically, in the second chapter I deal with content-based approaches to the problem of the nature of blame. These accounts of blame try to define the concept of blame in terms of which kind of mental state blame identifies with. In the third chapter, I introduce a different kind of account of blame. According to functional theories of blame, the instruments of “classic” philosophical analysis prove insufficient in dealing with the concept of blame. Thus, functional theories of blame address the question “What is blame?” by explaining what the function of blame is and what the point of our blaming practices is. According to functional theories of blame, whichever mental state serves that function counts as blame.

In the fourth and fifth of this dissertation, I put forward a theory of blame as signaling. According to this functional theory, the main point of blame is to signal the blamer's normative competence and normative commitments. In the fourth chapter, I present this theory and I defend it from objections. In the fifth and final chapter of this dissertation, I put the theory of blame as signaling to test in cases of non-moral blame. Can blame as signaling account for cases of non-moral blame, too? This question will also lead me to discuss a serious concern about non-moral blame. Is non-moral blame *proper* blame, or is it just a case of misidentified moral blame?



## CHAPTER ONE

### UNDERSTANDING BLAME

What does blame consist in, and what do we do when we blame each other? Is blame a belief, e.g., the belief that an agent has violated a norm or acted wrongly, or is it an emotional mental state such as resentment or indignation? Do we impose a form of sanction when we blame someone, or does blame serve a different kind of function? These are all fundamental questions for research on the nature of blame.

This dissertation will address all these important questions about blame in the following chapters. Before that, however, a series of preliminary questions will be asked. What is the object of the philosophical debate over blame? What kind of blame are ethicists interested in, and what should a theory of blame account for? The first chapter of this dissertation answers these questions, clarifying the object of my research. In §1, I distinguish different uses of the word “blame”; additionally, I motivate the choice of working directly on blame while not discussing problems related to free will and moral responsibility. In §2, I explain how most contemporary philosophers frame the debate on blame, distinguishing between theories on the nature of blame, and theories dealing with the normative structure of blame. Moreover, I motivate my choice of focusing on the problem of the nature of blame. In §3, I discuss an often-neglected topic in studies about blame: are praise and blame symmetrical concepts? In §4, I furtherly clarify the object of my research on blame by distinguishing between moral blame and non-moral blame. Finally, in §5, I propose a list of “data points”<sup>1</sup> of blame that a convincing theory of blame should account for. The five sections of this chapter lay the foundations for the rest of this dissertation, where I propose

---

<sup>1</sup> I borrow the expression “data points of blame” from a recent work by D. Shoemaker and M. Vargas (Shoemaker and Vargas 2021). By data points, I mean the features of blame that a satisfying theory of blame should be able to account for and explain.

a critical evaluation of the most prominent theories on blame, and I defend a particular approach to the study of blame.

## §1 Blame, Causation, and Blameworthiness

Throughout this dissertation, I will present several theories of blame. These theories try to correctly analyze the concept of blame, explaining how to discern between blame and different kinds of reactions to wrongdoing. If we want to understand the ongoing debate on blame, the first step will be tracing a distinction between two meanings of the term “blame”. In our ordinary language, it is not rare to encounter a usage of the word “blame” that is hardly ever discussed by philosophers. I am referring to situations in which we say that the lightning is “to blame” (or is blameworthy) for setting a barn on fire. A farmer, for example, could “blame” the lightning for this event. This seems to be just another way to say that the lightning caused the fire. Thus, as Chislenko notes, “the attribution of causation of an unfortunate event or state of affairs might be called *causal blame*” (Chislenko 2021, 347). In this causal meaning of the term, we may blame a cat for destroying our favorite vase of flowers, we may blame the dead battery of our car for not letting us show up at work on time, or we can blame young children for their actions – even if they do not understand the meaning of what they have done (Coates and Tognazzini 2013, 7-8). To my knowledge, the literature on causal blame is importantly lacking. One of the few reflections explicitly dedicated to causal blame comes from E. Beardsley. More than fifty years ago, she wrote “I consider the fact that we sometimes use ‘blame’ in the sense here discussed to be confusing and regrettable, and should like to recommend the adoption of some other locution for identifying a person causally responsible for an undesirable state of affairs” (Beardsley 1969, 38-39).<sup>2</sup>

---

<sup>2</sup> Similarly, Kenner noted that “in causal blame, all that we are doing is identifying the cause of some untoward event. [...] And this has nothing to do with moral disapproval” (Kenner 1967, 239).

Contemporary discussions of blame focus on another kind of blame. When we blame friends who do not show up on time, or when we blame dictators for their unjustifiable violence, or even when we blame ourselves for our faults, we are not merely establishing a causal link between agents and actions. We are doing, I would suggest, something more. As I will show in the next chapter, we can classify different theories on the nature of blame by specifying which elements have to be added to the causal link between an agent and an action. Moreover, as the former examples already suggest, the meaning of “blame” that interests most ethicists concerns our interpersonal or intrapersonal activities.<sup>3</sup> In this dissertation, I will be concerned about the definition and the analysis of this second meaning of the term “blame”. In other words, I will focus – as most philosophers do – on interpersonal blame.

Thus far, I have argued that the object of my research is interpersonal blame. What is needed, now, is a first, general definition of this kind of blame. This definition needs to be general enough to let us work towards a more qualified definition, but specific enough to be at least minimally informative. Fortunately, there are some available candidates for such a general definition. Consider these two passages from different authors.

Nearly all the leading theories of blame agree about at least this much: blame is a response to a person in light of his or her perceived norm violation, where the blamer takes that violated norm seriously (Shoemaker and Vargas 2021, 582).

Blame is a reaction to something of negative normative significance about someone or their behavior. A paradigm case, perhaps, would be when one person wrongs another, and the latter responds with resentment and a verbal rebuke, but of course

---

<sup>3</sup> Often, this kind of blame is labeled as “interpersonal blame” in order to distinguish it from causal blame (Coates and Tognazzini 2018). However, this kind of blame also applies to cases of self-blame, which are intrapersonal. Moreover, given the lack of interest philosophers demonstrate towards causal blame, it is usually unambiguous to refer to interpersonal blame as “blame” *simpliciter*.

we also blame others for their attitudes and characters (Coates and Tognazzini 2018, 1).

Albeit different, these general definitions suggest at least two things. First, as Björnsson correctly notes, blame has both a target and an object (Björnsson 2022): someone is blamed for something. In interpersonal blame, the targets are agents with relevant capacities – such as normative competence. The object of blame, however, can vary. We can blame agents for their actions, for their attitudes, for their omissions, and maybe even for their character. In this dissertation, I will focus on blame having actions as its target.<sup>4</sup> Moreover, even at the most general level, blame seems to be linked to actions of negative normative significance (or, as Shoemaker and Vargas claim, to “perceived norm violation”). Slightly more formally, we can say that blame is a reaction directed at X for Y, where Y is of negative normative significance. Naturally, such a definition is not very informative. However, this can be a starting point for our reflection on the nature of blame and on the norms governing our blaming practices.

Before proceeding, I would like to add a consideration about a concept that is strictly related to blame: the concept of blameworthiness. Blameworthiness is often tied to the concept of responsibility (Zimmerman 1988, McKenna 2012, Talbert 2022). Following Talbert, we can note that when X judges Y blameworthy for having performed action A, X is – at the very least – holding Y responsible for A (Talbert 2022, §1).<sup>5</sup> This point can be taken further by claiming that holding someone responsible involves responding to that person in ways that are made appropriate by the belief that the agent is responsible. And,

---

<sup>4</sup> In his book from 2006, G. Sher did an important work in clarifying how blame can have as its object the character of an agent (Sher 2006).

<sup>5</sup> Note that an agent can be responsible for something even without being blameworthy. The most obvious examples are instances of praiseworthy actions. Agents may be responsible for praiseworthy actions, and this would imply that they are not blameworthy for them. See §3 of this dissertation for considerations about praise and praiseworthiness. See also McKenna 2012, 16-17 for his interesting suggestion about “neutral” exercise of agency, e.g., agents who are responsible for an action but are neither blameworthy nor praiseworthy for having performed that action.



among these ways, blame is often an appropriate response to wrongdoing. As I have already stated in the Introduction to this dissertation, however, my aim is to provide an analysis of the concept of blame, rather than that of blameworthiness. Thus, I prefer to remain neutral on the debate about the conditions of blameworthiness.<sup>6</sup> Thus, following Graham, I treat the concept of blameworthiness in a minimal way (Graham 2014). According to this minimal analysis, an agent A is blameworthy for X-ing just in case A is worthy of blame for X-ing (Graham 2014). In addition, A is worthy of blame if and only if it would be appropriate for A to be blamed for X-ing. Albeit plausible, this analysis runs the risk of being trivial. Up to now, I still have not said what exactly means for A to be blamed for X-ing. This is precisely what is at stake in this dissertation. What is blame, and how can we discern blame from other responses to wrongdoing?

In the next section, I will frame the contemporary debate on blame, and I will motivate my choice of focusing on the concept of blame and the practice of blaming each other rather than contributing to the debate on moral responsibility and free will.

## **§2 Framing the Debate on Blame**

In the introduction to this chapter, I have cursorily used the expression “the nature of blame”. What does this mean? In this section, I clarify my usage of this expression, showing that most authors refer to the term “nature of blame” to frame a set of problems. Understanding the set of problems identified by this expression is of fundamental importance for the structure of my dissertation. Indeed, my hope is that this dissertation will contribute to the debate on the nature of blame. Moreover, this section helps the reader understand how most authors frame the ongoing debate on blame.

---

<sup>6</sup> However, for the sake of the argument, I will assume that at least some agents are actually blameworthy. In other words, I will not consider discuss against the concept of blameworthiness (Rosen 2004).

I want to start by drawing an important distinction. One can distinguish studies on blame into two fields. In turn, these two fields are constituted by different questions. The first field is often labeled as the debate on the nature of blame (Coates and Tognazzini 2013; 2018). The set of questions that constitutes this field is: what exactly is blame? Is it possible to use the traditional instruments of conceptual analysis in order to define the necessary and sufficient conditions of blame? Is blame an activity, a belief, an emotion, or some other mental state (or a set of mental states)? Moreover, the debate on the nature of blame also includes questions about the function of blame: what is the function of blame? Is it possible to explain the concept of blame and our blaming practices by identifying the main function of blame? The second set of questions focuses on the norms governing our blaming practices: are there conditions on the appropriateness of blame? And, if so, what are these conditions? At least intuitively, it seems that we can blame agents in an apt or inapt way. What are the criteria for distinguishing apt blame from inapt blame? The first field that I have identified is often referred to as the problem of the nature of blame. The second field, on the other hand, is often referred to as the debate surrounding the ethics of blame.<sup>7</sup> Some questions about the nature of blame may overlap with problems related to the nature of blame – and vice versa. For clarity, in what follows I try to keep questions surrounding the nature of blame separated from those surrounding the ethics of blame.

As I have already said, the expression “nature of blame” frames a set of problems related to the correct way to distinguish blame from other interpersonal reactions. If we have a criterion to identify blame (as, for example, a set of necessary and sufficient conditions), then we can understand when an agent is blaming another, distinguishing blame from other kinds of reactions. Among the different theories on the nature of blame, it is possible to

---

<sup>7</sup> “Ethics of blame” is an expression used by T.M. Scanlon to identify the set of normative questions surrounding blame. In his words, the ethics of blame concerns facts “about who can be blamed, who has standing to blame, and why we should blame—why blame is not an attitude we would do better to avoid” (Scanlon 2008, 123).

identify two different methodological approaches. According to some, the most appropriate way to answer the questions surrounding the nature of blame is to describe blame as a mental state. The project of identifying blame with a mental state has been recently clearly expressed by P. Sliwa. She correctly argues that “the received wisdom is to think of blame as a psychological phenomenon: a mental state. Consequently, a central line of philosophical inquiry has focused on the question what kind of mental state it is: an emotion, a desire, a judgment, or some combination thereof” (Sliwa 2021, 200). In this dissertation, I will frequently refer to theories of blame that identify blame with a kind of mental state as content-based theories.<sup>8</sup> As I will show in the next chapter, content-based theories can be distinguished into at least three different categories: cognitivist accounts, emotional accounts, and conative accounts. Cognitivist accounts of blame maintain that blame is a cognitive mental state such as a judgment or a belief; emotional accounts of blame argue in favor of the identification of blame with an emotion; conative accounts state that the correct definition of blame involves conative mental states such as desires and intentions (Coates and Tognazzini 2013).

Content-based theories on the nature of blame, however, are just part of the picture. Very recently, an increasing number of philosophers expressed their skepticism towards the project of analyzing blame by identifying a mental state (or a set of mental state). This skepticism is well described by M. Nussbaum. According to her, “insofar as all these fine philosophers [authors who try to identify blame with a mental state] are pursuing a single essence, they appear to be pursuing a will-o’-the-wisp” (Nussbaum 2016, 260). By the end of the next chapter, the reader will find this skeptical stance well motivated – at least, I hope so. If content-based accounts of blame fail in their task of answering the questions surrounding

---

<sup>8</sup> D. Shoemaker and M. Vargas use this term to refer to the group of theories that identify blame with a mental state. Even if I do not completely agree with the taxonomy they propose, I borrow the expression “content-based theories” from their recent paper *Moral Torch Fishing* (Shoemaker and Vargas, 2021).

the nature of blame, then how should we proceed? In the past fifteen years, some authors started vindicating a different approach to the study of the nature of blame. According to this recent approach, questions surrounding the nature of blame should be answered by taking a functionalist stance. According to functional theories of blame, it is not important to identify a single mental state like a belief or an emotion. According to functionalist approaches, blame is just whatever mental state serves a particular function.<sup>9</sup> Quite obviously, every functional account of blame identifies a different function of blame. According to some authors, for example, the function of blame is to protest (Talbert 2012; Smith 2013), while others maintain that the function of blame is to bring increased alignment of the moral understanding of wronged and wrongdoer (Fricker 2016).

Both content-based and functional accounts of blame attempt to find a plausible solution to the questions on the nature of blame. However, the nature of blame is just part of the picture of the ongoing debate on blame. I have already mentioned the term “ethics of blame”. In the remaining of this section, my aim is to shed light on the debate surrounding the normative structure of blame. M. Fricker provides a good starting point for my discussion of the ethics of blame: she argues that just like many other human practices, “blame is susceptible to the vices of being done from the wrong sort of motive, in the wrong degree, in the wrong way, or with the wrong sort of object” (Fricker 2016, 168). The debate on the ethics of blame, in other words, does not focus on what blame is, but on when blame is appropriate. In order to understand when blame is appropriate, we can distinguish at least three different sets of considerations (Coates and Tognazzini 2021).

First, we need to take into account facts about the agent being blamed. As I have noted in the first section of this chapter, we can distinguish two uses of the term “blame”. I

---

<sup>9</sup> As Polger puts it, then, blame is more like a mousetrap than a diamond (Polger 2019). For a discussion, see my Introduction of Chapter Three of this dissertation.

have argued that what concerns me in this dissertation is interpersonal blame, and not causal blame. Blaming dead car batteries or lightning in the interpersonal interpretation of “blame” would not only be pointless, but also inappropriate.<sup>10</sup> This suggests that, when we blame X, our blame is appropriate only if X possesses certain properties. Now, the exact list of properties an agent must have in order to be appropriately blamed is matter of discussion among philosophers.<sup>11</sup> Generally, it is safe to say that - at the very least - an agent must be normatively competent: she must be capable of reflecting upon, reasoning about, and executing a decision about how to behave (Coates and Tognazzini, 2021). Normative competence, however, could be insufficient. At this point, discussions about free will and moral responsibility become relevant. If no one has the possibility to do otherwise, then the concept of blameworthiness may be at threat. How could blame be appropriate if its target has acted in the only way he could have acted? J.M. Fischer and M. Ravizza, however, convincingly showed that the possibility to do otherwise could not be a necessary component of the appropriateness of blame (Fischer and Ravizza 1998). Moreover – as I will explain more in length in the next chapter – P.F. Strawson argued that our blaming reactions do not need an “external” (read: metaphysical) justification (Strawson 2008, 25).

As I have just shown, we could ask various questions about the agent being blamed: was the agent free to act? Is the agent normatively competent, or is she incapable of reflecting on her own decisions and act upon them? These questions are surely important, and they have always been strictly related to the debate on free will and moral responsibility. However, they do not represent everything that could be said about the appropriateness of blame.

---

<sup>10</sup> Note that I use “inappropriate” and not “(morally) wrong” here. This lets me discuss the ethics of blame in a broad sense. According to a narrower interpretation, the ethics of blame does not deal with “inappropriate” blame, but on “wrong” blame, where this means that blame has problematic moral implications. Certainly, I think that misplaced blame could have moral implications (think of when we harshly blame someone who has a valid excuse), but this does not mean that we should limit the domain of the ethics of blame to the morally problematic feature of blame. No doubt moral considerations are important, but I find it better to favor a broader interpretation of the ethics of blame and also include questions about blame appropriateness (or aptness).

<sup>11</sup> In his 2015, D. Shoemaker provides an insightful discussion on this topic.

There are at least two other important sets of problems worth discussing. Blame could not only be rendered inappropriate by facts about the blamed agent, but also by facts about the blaming interaction, and by facts about the blamer.

Consider this example of blaming interaction. Suppose that I have just received a phone call from my friend. He has asked me to meet him at the local cafeteria, and I have agreed to his request. We are supposed to meet at six o'clock. However, on his way to the cafeteria, my friend realized that he left his wallet at home. Luckily, he lives near the cafeteria, so he is just five minutes late. So, I have to wait five minutes seated alone at a table. As soon as he arrives, I start yelling at him "how could you be late! This is the last time you and I hang out together!". To add some details to this example, imagine that my friend being late has no important consequences – maybe, we are just meeting to chat and spend some time together. Moreover, my friend has always been on time on other occasions. Under these conditions, my blaming reaction would certainly be disproportionate – or, better, it would be inappropriate because of its being disproportionate. Thus, this suggests that proportionality is important in assessing the appropriateness of blame.<sup>12</sup>

Another fact about the blaming interaction that could make blame inappropriate is the foreseeable outcome of the blaming activity. This point applies mostly to overt blame, e.g., blame that is outwardly expressed.<sup>13</sup> Sometimes, an agent satisfies all the applicable conditions for blameworthiness but, despite this, blaming that agent would bring about terrible consequences. Undoubtedly, being on the receiving end of a blaming interaction may be hurtful. Being blamed may have psychological effects on the blamed agent that may render blame inappropriate. Once a prospective blamer has considered the possible effects of blame,

---

<sup>12</sup> See Smith 2007 for a more refined case which involves an agent who has already acknowledged her mistakes: "If someone has an objectionable attitude toward me, for example, but is already reproaching herself for it and making efforts to change, then I may judge that I have no reason to adopt or express any blaming attitudes toward her at all" (Smith 2007, 482).

<sup>13</sup> See §5 of this chapter.

he or she may ponder whether blame is the appropriate reaction. Sometimes, another kind of reaction would be more appropriate. For example, we can just point out the fact that the agent has performed a wrong action, thus avoiding the characteristic sting of blame.<sup>14</sup>

This leaves us with facts about the blamer. I find this set of problems particularly interesting and worth discussing, albeit in this dissertation I mostly focus on the debate surrounding the nature of blame. At least two facts about the blamer are worth mentioning here. The first is eminently epistemic, while the second concerns the blamer's standing. Let us start from the epistemic concern. As P. Todd notes, blame is appropriate only if the blamer is justified in believing that the agent who is being blamed is blameworthy (Todd 2019, 4). This means that blame raises some concerns about the blamer's warrant in believing that the target of blame is actually blameworthy. As Coates and Tognazzini note, "The realm of interpersonal blame is not perfectly analogous to the realm of legal responsibility, of course, so "beyond a reasonable doubt" may be too demanding a requirement, but nevertheless there is some epistemic standard that must be met before blame is appropriate, even if the potential target of blame is in fact blameworthy" (Coates and Tognazzini 2021). If, for example, X has just heard some rumors about Y, blaming Y would be inappropriate for X (but not for someone else).

Finally, blame can be made inappropriate by the lack of standing of the blamer. In order to introduce this point, consider this example. Every year, a friend of yours forgets your birthday. You do not give particular weight to her forgetfulness, even though you have never forgotten her birthday. One year, however, you forget her birthday. In response, your friend openly blames you. On this occasion, you may think that there is something off in this blaming interaction; namely, you may think that this blaming interaction is inappropriate. It seems that the most natural way to address your friend is by saying "Who are you to blame

---

<sup>14</sup> For an accurate analysis of the "stinginess" of blame, see Hieronymi 2004.

me?”). How can we account for this phenomenon? First, we can note that your friend is acting as a hypocrite (Szabados and Soifer 2004). Now, that would not be sufficient, for it does not explain why hypocrisy renders blame inappropriate. There is a growing literature that addresses the problem of hypocrisy in relation to blame (Wallace 2010; Radzik 2011; Bell 2012a; Herstein 2017; Fritz and Miller 2018; King 2019). Since the ethics of blame will not be central in this dissertation, I do not intend to discuss hypocrisy at length. However, I think that Fritz and Miller advanced our understanding of the problem of standing to blame and hypocrisy by claiming that the hypocritical blamer “forfeit the right to blame others” for violations of the norms she uses to violate herself (Fritz and Miller 2018, 125). This way, it seems that we can account for our intuition about the inappropriateness of hypocritical blame. Even if X is blameworthy for having violated the norm N, Y could have forfeited his right to blame X by having repeatedly violated N herself. This means that, while others could still appropriately blame X, Y could not.<sup>15</sup>

In this section, I have offered an overview of the contemporary debate on blame. In particular, I have shown that philosophers distinguish between the problem of the nature of blame and the problem of the ethics of blame. Both content-based theories and functional theories attempt to answer the questions surrounding the problem of the nature of blame: what is blame, and how can we define it? How can we distinguish blame from other different reactions? Content-based accounts, which I will analyze in Chapter Two, identify blame with a kind of mental state (cognitive, emotional or conative). Functional accounts, which I will analyze in Chapter Three, maintain that what distinguishes blame from other reactions is the function of blame. This way, whatever mental state serves that function is considered blame. In the last part of this section, I have argued that the ethics of blame includes three different

---

<sup>15</sup> I believe Fritz and Miller nicely capture our intuition about hypocrisy and the standing to blame. However, I do not want to suggest the idea that their proposal is universally accepted. On the contrary, some philosophers are even skeptical about the plausibility of the concept of standing to blame. For a skeptical approach, see Bell 2012a and King 2019).



sets of questions concerning the norm of appropriateness of blame. First, we find questions about the agent who is being blamed: is the agent normatively competent? Was the agent free to act? Second, we have questions about the blaming interaction: is blame proportionate to the wrong action? Will blame have extremely negative consequences? Finally, we find questions about the prospective blamer: is the blamer justified in thinking that the person being blamed is actually blameworthy? Does the blamer have the appropriate standing to blame?

Before moving to the next section, I want to spend some words motivating my choice of focusing my effort on the problem of the nature of blame. First and foremost, a quick review of the literature on the nature of blame reveals that the problem is far from being solved. Year after year, many authors attempted to answer the questions surrounding the nature of blame. While every account has its own advantages over others, I am not convinced by most of them. Thus, the need to defend an account of the nature of blame that I find more convincing than others. Second, while in principle the nature of blame and the ethics of blame can be considered as separate problems, I am inclined to think that a convincing account of the nature of blame can help us answering at least some questions surrounding the ethics of blame. Thus, I find it important to focus on the problem of the nature of blame.

In the next section, I would like to consider the relationship between blame and a concept often mentioned alongside it: praise.

### **§3 Blame and Praise**

Before proceeding, I would like to briefly focus on an often-neglected topic in the literature on blame. It is not until recently that philosophers started considering blame directly, and not as a by-product of discussions on free will and moral responsibility. In the past three decades, however, works on blame as an interesting phenomenon in moral psychology and more broadly in ethics have been proliferating. It is common to think of praise as a concept

germane to the concept of blame (Coates and Tognazzini 2013; Talbert 2022; Stout 2020). More specifically, praise is often considered as the “positive counterpart” of blame. While blame “is a response that may follow on the judgment that a person is morally responsible for behavior that is wrong or bad, praise is a response that may follow on the judgment that a person is morally responsible for behavior that is right or good” (Talbert 2022, §1). Most authors leave this point as an assumption, and do not delve deeper into this matter. In other words, the conceptual symmetry between praise and blame is just assumed. As it almost always happens in philosophy, however, assumptions can (or often, should) be challenged. In this section, my aim is challenging the assumption that blame and praise are symmetrical concepts, and that our praising practices mirror our blaming practices. I call this the “Symmetry Assumption”.

I would like to begin by saying that I do not have a knock-down argument in favor of the asymmetry of the concepts of blame and praise. More modestly, my aim here is to challenge the Symmetry Assumption by fostering some doubts over it. First, I shall clarify an important point. From a descriptive point of view, it is already clear that there are at least two asymmetries between blame and praise. First, it is already clear that the attention given by philosophers to the concept of blame greatly exceeds the attention they have given to the concept of praise (Coates and Tognazzini 2013).<sup>16</sup> Second, from a linguistic point of view, Watson points out that:

We seem to have a richer vocabulary of blame than praise. This slant is not due solely to mean-spiritedness. At least part of the explanation is that blaming tends to be a much more serious affair: reputation, liberty and even life can be at stake, and understandably we are more concerned with the conditions of adverse treatment than with those of favourable treatment (Watson 2004, 283).

---

<sup>16</sup> From this point of view, this dissertation does not hell in reverting the imbalance.

Watson does not only mention the fact that we have a richer vocabulary of blame than praise. He goes on to argue that a partial explanation for this fact is due to what is at stake in our blaming interactions. Surely, it can be conceded that blame is sometimes linked to harsh treatment.<sup>17</sup> However, not praising a praiseworthy agent could represent a great form of unfairness, or even injustice. So, despite the fact that we can agree with Watson and arguing that we are more concerned with the conditions of adverse treatment than with those of favourable treatment, we should also provide an analysis of praise and an explanation of our praising practices.

The two points above merely identify a descriptive asymmetry between praise and blame. In the remaining of this section, I want to challenge the Symmetry Assumption by showing that we can identify asymmetries even from a conceptual point of view. First, consider this minor point. Most philosophers agree on the fact that blame can be kept private rather than being expressed (Wallace 1994; Scanlon 2008; Smith 2013). Blame can be kept private for several reasons. Imagine this scenario. I am alone at my house writing my dissertation. While taking a break, I decide to read an online newspaper. When I read that a politician was found guilty of corruption, I blame that politician. However, being alone at my house, I cannot communicate my blame to anyone. Certainly, I cannot express my blame in front of the politician. But I cannot even express my blame to my partner or my friends, since I am alone. In §5, I provide a longer discussion of cases of private blame. However, even this simple example should be enough to grasp the concept of private blame. It seems much less plausible to think that praise can be kept private in the same way. Consider this other scenario, suggested by Coates and Tognazzini in their 2013:

---

<sup>17</sup> As I will often argue throughout the dissertation, I do not think that blame constitutes a form of harsh treatment. More radically, I think that often blame does not constitute a form of treatment at all.

You might discover to your horror that your spouse continues to blame you for something you did several years ago, but it seems at best awkward to say that you might discover that your spouse has been praising you for several years without anyone's knowing about it (Coates and Tognazzini 2013, 5).

One could object that this asymmetry between praise and blame is only apparent, and it could easily be solved by distinguishing between the concept of praise and our actual praising practices. It may well be that our actual practices suggest an asymmetry between praise and blame, e.g., that blame can be kept private while praise cannot. But this does not mean that, at a conceptual level, praise and blame are not symmetrical. Thus, the Symmetry Assumption may still hold.

The conceptual point that I want to raise against the Symmetry Assumption (although, as I have already written, I do not take it to be a knock-down argument), is suggested by A. Eshleman (Eshleman 2014). Consider this familiar example. As bystanders, we would probably blame a driver who has just ignored the red light. That driver showed a lack of concern for the safety of both pedestrians and other drivers. This suggest – as I have written in §1 – that we blame agents for their conduct when this conduct is perceived as a violation of a norm. In other words, we sometimes blame agents when their actions are subpar given a normative standard. If praise were the “positive counterpart” of blame, then we would praise agents whose actions successfully meet a normative standard. But this does not seem to be true. Consider a careful driver who always stops at the red light. As bystanders, would we praise her every time she stops at the red light? This seems counterintuitive. It could be objected, however, that this is only part of the picture, and it does not pose any threat to the Symmetry Assumption. Sometimes, we actually praise agents who merely meet their duty, as we may say. Consider this other example. During a written examination, the professor has to leave the room for a few minutes. As Eshleman writes, “imagine your

reaction upon learning of a student who refrained from cheating while all those around her were cheating” (Eshleman 2014, 222). It could be argued that we would praise the student for his conduct. The student, however, has merely met a normative standard: no one should cheat during an examination. Thus, while we blame the other students for having failed to meet a normative standard, we praise the honest student for having met the same standard. If this is so, then the Symmetry Assumption still holds. We should, however, carefully consider what is the target of both our blaming activity and our praising activity. Upon close inspection, it could be argued that while we blame the cheating students, we are blaming them for their action, e.g., cheating. Conversely, when we praise the honest student, we are praising her because in not cheating she revealed her integrity and her honesty, and not merely because she has not cheated. If this is so, then, my criticism against the Symmetry Assumption still holds, for the objects of blame and praise in the above example are different. Thus, the example does not show that we blame agents who fail to meet a standard, while we praise agents who meet a normative standard. Upon reflection, the example suggests that the objects of praise and blame are different. And this is precisely where a conceptual asymmetry between praise and blame can be found: where the object of blame is often a subpar action performed by an agent, the object of praise usually is an admirable quality of character.

Again, my aim here was not to provide a definitive argument against the Symmetry Assumption. Instead, I wanted to reflect on an often-neglected topic in the literature on blame. My hope is that, in the near future, more and more philosophers will stop assuming the symmetry between praise and blame as a *datum* that is not worth discussing. In the next section, I will undertake another topic in the blaming literature that has not received much attention – albeit more than praise.

#### §4 Moral Blame and Non-moral Blame

Let us now turn back to the concept of blame and to our blaming practices. It is not hard to see that blame is a familiar part of our everyday life. We often react to agents' faults by blaming them: as I will argue in the next sections, the variety of blame may constitute a problem for ethicists who want to offer an analysis of blame's necessary and sufficient conditions. The project of offering a detailed analysis of the concept of blame and a convincing explanation of our blaming practices is furtherly complicated by the following consideration: at least intuitively, it seems that blame is not confined solely to the moral realm. If we reflect on our everyday experience, we find that we do not blame agents only for their moral failures. Does this mean that a convincing theory of blame should take into account cases of non-moral blame? Or, upon reflection, is non-moral blame just misidentified moral blame? In this section, I provide an answer to these questions. In particular, I argue that a convincing theory of blame should also extend to non-moral domains, and that there are cases of authentic non-moral blame, e.g., blame that is not just misidentified moral blame.

Let us begin by mentioning some cases of non-moral blame. Imagine your friends inviting you over in order to watch a football competition on the TV. You are not a huge football fan, but you enjoy spending time with your friends, so you decide to join them. Part of your enjoyment in spending time with them comes from the fact that they are very passionate about football, and they always openly express their opinions on players and the referee. During their game, one of their favorite players misses a penalty kick: the team your friends are rooting for, as a result, loses the game. Your friends openly blame that player, insisting on the fact that his technique was poor, and he just played badly. Being a good football player, however, does not seem to be a moral standard. Thus, this seems a case of

non-moral blame: your friends are blaming an agent for an action he performed, where this action is not morally sub-standard.

Consider this another – perhaps less familiar – example. Chess lovers should remember a well-known event in the recent history of chess. In 2014, Magnus Carlsen faced Viswanathan Anand for the World Champion title. During game 6, Carlsen performs a big blunder. In the video recording of this episode, we can clearly see his reaction upon realizing that he has conceded a great advantage to his opponent. His visible relief is even greater when Viswanathan Anand misses his opportunity to win the game.<sup>18</sup> For people who do not play chess, Carlsen’s blunder may not seem a great error. However, it is plausible to think that upon realizing his blunder, Carlsen feels the urge to kick himself for the mistake he has just performed on the chessboard. Carlsen is a well-known skilled chess player, and he has trained appropriately to face his opponent. Carlsen could have easily avoided his mistake, and he even had the advantage prior to his bad move. His failure was not due to unfortunate circumstances, nor had he failed to prepare adequately for the match. Moreover, winning the game was in his own self-interest. Carlsen realized that the mistake was his own fault: he fell short of certain chess-related standards. Upon realizing that, he blames himself for his blunder.

Both examples suggest that we occasionally blame agents for their non-moral faults. Thus, it seems that both our interpersonal and intrapersonal practices include cases of non-moral blame. However, the contemporary literature on blame rarely focuses on non-moral blame. Many authors acknowledge the distinction between moral and non-moral blame (Press 1969; Eaton 2008; Rorty 2010; Shoemaker 2015; Luthra 2016; Peels 2016; Archer and Ware 2018; Björnsson 2017), but only a few takes non-moral blame as the primary object of

---

<sup>18</sup> The interested reader can find the notation of the whole match following this link: <https://www.chess.com/article/view/the-7-most-shocking-world-championship-blunders>.

their research (Tollefsen 2017; Brown 2018; Nelkin 2020). The choice of focusing on the concept of moral blame and of our practices of morally blaming each other should not surprise us. Moral wrongs are often (perhaps, always) more severe than non-moral ones. Surely, we want a theory that is able to explain what it means to blame a child abuser or a violent dictator, rather than a theory that is able to explain why and how we blame football players for their sloppy plays. However, I think that a convincing theory of blame should be able to explain both moral and non-moral blame. Such a theory would have a greater explanatory power, being able to address cases of blame both in the moral domain and other normative domains. Moreover, as the above examples suggest, instances of non-moral blame are already present both in our interpersonal practices and in our intrapersonal practices. Despite this last consideration, some authors have argued in favor of a different approach to the study of blame. According to this approach, blame is essentially moral. In other words, at least from a conceptual point of view, non-moral blame is impossible. Even if I do not find this thesis to be true, I find it important to discuss it. So, let me take into account what we may name “the Conceptual Challenge” to non-moral blame.

At least two prominent authors suggest the idea that non-moral blame is impossible. B. Williams writes that “Blame is the characteristic reaction of the morality system” (Williams 1985, 177), and he later adds that “blame always tends to share the particularized, practical character of moral obligation in the technical sense. Its negative reaction is focused closely on an action or omission, and this is what is blamed” (Williams 1985, 177). Williams here suggests that the link between blame and the moral domain is inescapable. This tie is strengthened by the fact that blame and moral obligations share the same “practical” character. Compare William’s thesis with this passage from S. Darwall: “Moral obligation really is conceptually related to standards of minimally decent conduct that moral agents are accountable for complying with. And the forms of moral accountability—blame, guilt,



indignation, punishment, and so on—really do imply that agents have reasons (indeed, conclusive reasons) to do what they are morally obligated and accountable for doing” (Darwall 2006, 94). Here, Darwall states that blame is a form of “moral accountability”. What if Williams and Darwall are right, and blame is inextricably tied with the moral domain? If this were true, then a convincing theory of blame should exclude discussing non-moral blame, since all blame would be moral blame.

In order to reject this thesis, I have at least two observations to express. First, it should be noted that making a conceptual point may not be enough. As I have argued, it seems that our actual practices include cases of non-moral blame. Recall my previous examples of cases of non-moral blame. If Darwall and Williams are correct, our intuitions about those examples are wrong. Your friends do not blame the football player who missed his opportunity to score for his team, and Carlsen does not blame himself for his chess mistake. Let us assume that Williams and Darwall are right, and that my friends and Carlsen are not blaming anyone. Then, what are my friends and Carlsen doing? It seems to me that the most plausible model to explain and understand their practices is referring to the concept of blame. Then, why not just distinguish moral blame from non-moral blame, and concede that we actually blame agents for their non-moral faults? Of course, this is not enough to dismiss Williams’s and Darwall’s point of view. However, there is another important point to be made.

The conceptual argument against non-moral blame commits to an emotional account of blame. As I will show in the next chapter, however, emotional accounts of blame cannot account for the variety of blame that I have described throughout this chapter.<sup>19</sup> Thus, if one wants to accept the conceptual point against non-moral blame, one also must accept a substantive theory about the nature of blame. According to this theory, blame is primarily an

---

<sup>19</sup> The variety of blame will be entirely manifest in the next section, §5.

emotion, and there are no instances of dispassionate blame. As I will argue in the next chapter, however, this position can be attacked. But there is more. It is not sufficient to accept an emotional account of blame in order to agree with Williams and Darwall on the fact that non-moral blame is conceptually impossible. In fact, one would also have to come up with an emotional account of blame that is not centered on the reactive attitudes, since it is possible to show that our reactive attitudes can be extended to other normative domains than morality.<sup>20</sup> Thus, it seems that the Conceptual Challenge against non-moral blame suffers from at least these two problems: first, it does not clearly explain our intuitions about cases of alleged non-moral blame; second, at least in the Williams-Darwall's version of it, it hinges on the plausibility of an emotional account of blame.

I think that our intuitions about the examples that I have shown in the beginning of this section are correct. I have highlighted that, from a conceptual point of view, non-moral blame should also be accounted for by a convincing theory on the nature of blame. In the fifth chapter of my dissertation, I will discuss another challenge against non-moral blame. According to this challenge, the problems linked to non-moral blame do not arise from a conceptual point of view. P.E. Milam and B. Matheson have recently argued that proponents of non-moral blame often fall into the “Moralizing Trap”, which consists in the fallacy of misidentifying cases of moral blame as instances of non-moral blame. According to this view, most cases of alleged non-moral blame are, upon reflection, cases of moral blame (Matheson and Milam 2021). I find their view interesting and their argument well-structured. Thus, I consider their case against non-moral blame to be a real challenge to defenders of non-moral blame. Since I consider myself among defenders of the plausibility of non-moral blame, I shall find a solution to the problems they have raised. In order to give a plausible answer to their questions on non-moral blame, however, I first need to defend a substantial account of

---

<sup>20</sup> For reference, in her 2017 D. Tollefsen showed that the reactive attitude account of blame can be extended to the epistemic domain.

blame. Because of this, I prefer delaying the discussion of the “Moralizing Trap” to the last chapter of my dissertation.

Despite the intuitive link between blame and moral faults, in this section I have shown that we sometimes blame agents for their non-moral faults. Then, I have presented a challenge to the concept of non-moral blame. According to this challenge, all blame is moral blame: there is no such thing as non-moral blame. I think most accounts are well equipped to find a solution to this challenge. However, proponents of non-moral blame face another problem. What if we find that what we call “non-moral blame” is nothing but misidentified moral blame? In other words, what if – upon close inspection – cases of alleged non-moral blame are cases of moral blame? I will return to these questions after having defended a substantial theory of the nature of blame. In the next section, I turn to a fundamental question: what should a theory of blame account for?

### **§5 What should a Theory of Blame Account for?**

In this section, I will present what I consider to be the *desiderata* of a compelling theory of blame. This means that a convincing theory of blame should be able to account for every element in this list. As I will argue in the following chapters of this dissertation, most theories on the nature of blame fail to convincingly account for all the “data points” of blame (Shoemaker and Vargas 2021).<sup>21</sup> Before proceeding to the list of data points, I want to add a preliminary remark. I am aware that some theories on the nature of blame do not recognize some elements of the list as *desiderata*. For example, emotional theories obviously reject dispassionate blame, believing that one cannot properly speak of blame in cases where the blamer is not in a certain emotional mental state. I will address these kinds of issues related to data points as I present them. The aim of this section is to highlight a difficulty linked to

---

<sup>21</sup> The “data points” of blame are the elements of our blaming interactions that a convincing theory of blame should be able to explain.

the study of blame that the vast majority of philosophers recognize. I am referring to the complex variety of our blaming interactions (Scanlon 2008; Coates 2012; Fricker 2016; Sliwa 2021; Coates and Tognazzini 2021): if a theory of blame fails to account for some elements on this list of data points, it could mean that such theory is inappropriate to explain the phenomenon of blame. At best, this could mean that the theory needs to be revised; at worst, it could mean that the theory should be abandoned in favor of another approach.

Here is the list of blame's data points that a convincing theory of blame should be able to account for.

- Blame involves more than the mere belief that the norm violator has acted wrongly (Sher 2006, 6); at the same time, as Wallace notes, “the essence of blame equally cannot be understood in terms of the actions that we might perform in response to an episode of wrongdoing or disregard. In particular, we can blame someone without undertaking to punish them” (Wallace 2011, 348). Recognizing that blame encompasses more than simply believing that an agent has acted wrongly helps us distinguish instances of blame from the belief that an agent is blameworthy.<sup>22</sup> If blame were not at all different from the mere belief that an agent acted wrongly, then it would not be possible to separate the concept of blame from that of blameworthiness. Thus, every instance of this type of belief would constitute blame. However, simply believing that you have wronged me does not amount to blame. The person who has been wronged may even be happy or acquire some kind of benefit from being wronged (McKenna 2012). Therefore, it seems plausible to believe that blame includes additional elements compared to the belief that an agent has acted wrongly. As I will show in Chapters 2 and 3, what distinguishes blame from this belief is the subject of extensive discussion in the literature on blame. Finally, blame does not necessarily amount to a form of punishment. Not only the concepts of punishment and blame are distinguishable

---

<sup>22</sup> More on this in §1, Chapter Two.

(Shoemaker 2013); it is also possible to imagine cases in which someone blames an agent without expressing her blame to anyone (as in cases of private blame – more on this later).

- Directed dyadic overt blame: I take it to be the least controversial form of blame that a plausible theory of blame should account for. This blaming interaction involves two parties (hence, the "dyadic" attribute), wherein one party overtly blames the other for something. The blamer overtly and directly communicates her blame to the blamed agent through gestures, words, and even omissions. At least implicitly, many theorists take this form of blame to be paradigmatic.<sup>23</sup> M. Fricker explicitly defends the thesis that this form of blame is paradigmatic and even explanatorily prior to other forms of blame. As I will argue in §3, Chapter 3, I disagree with this thesis. However, directed dyadic overt blame is undoubtedly an important blaming interaction that should be explained by a theory of the nature of blame.

- Third party overt blame: we can blame someone even without being face-to-face with the person being blamed. Although I do not have empirical data to support this statement, I suppose it is plausible to believe that this type of blame is even more widespread than the one described earlier. There are several reasons in favor of this supposition. On the one hand, it is not always possible to blame someone directly because the blamed agent could simply be absent. She might be in another room, another city, or even on the other side of the world, for all that matters. Moreover, there are prudential reasons for not blaming someone directly. If I directly blame a dangerous person, I could get myself into trouble. Overtly and directly blaming an agent could provoke a reaction from the person being blamed. And sometimes, it is preferable – for prudential reasons – not to provoke any reaction at all. Instead of directly addressing the target of our blame, we could blame her

---

<sup>23</sup> However, I want to point out that there are at least two interesting theories of blame that identify private blame as explanatory prior to directed dyadic overt blame (Arpaly 2006, 9; Carlsson 2017).

with our friends, co-workers, family or bystanders: this is a widespread phenomenon. I believe that a good theory of blame should be able to account for this form of indirect blame as well.

- **Private Blame:** We can blame someone both by outwardly expressing our blame (through gestures or words) and by keeping the blame private. As I briefly discussed in §3, this aspect seems to distinguish blame from praise: while it seems plausible to privately blame someone, the idea of private praise does not seem equally plausible. As with indirect blame, there may also be prudential reasons for keeping the blame private. For example, this is true when we know – or can reasonably foresee – that we would suffer negative consequences if we were to express our blame publicly. As I will show in the next chapter, accounting for the possibility of private blame constitutes a serious challenge for some theories of the nature of blame. This challenging aspect can be explained by considering that most theories of the nature of blame begin by analyzing the concept of blame – and our blaming interactions – starting from overt blame. When moving from the concept of overt blame to that of private blame, however, some theories need to be revised, as some considerations on overt blame do not apply to private blame. In the second and third chapters of this dissertation, the reader will see how private blame constitutes a challenge to theories on the nature of blame. For now, it is important to note that even theories that face problems in accounting for private blame do not exclude this type of blame from the list of data points of blame (Talbert 2012; Smith 2013). According to my interpretation, this is a good sign of the plausibility of considering private blame as a data point.

- **Self-Blame:** It does not seem particularly controversial to consider self-blame among the data points of blame. After all, self-blame can be defined as a case of blame in which the blamer coincides with the blamed. In a more formal language, in private blame X blames agent Y, where X and Y are the same person. Individuals can blame themselves both

overtly and privately.<sup>24</sup> I see no particular reasons not to include self-blame in the phenomena that a good theory of the nature of blame should be able to explain. For this reason, self-blame fits without much trouble into the data points of blame.<sup>25</sup>

- Blaming the distant (both in time and in space): I have already noted how – in cases of indirect or third-party blame – one can blame a person even in her absence. This typically occurs in cases of third-party or "gossipy" blame,<sup>26</sup> but it is not the only case. After all, we can also privately blame someone in her absence. This suggests that we can blame agents if they are not present. This idea seems intuitive and plausible, and most theories of blame can explain the phenomenon of blaming the absent. A more controversial case, on the other hand, is that of blame directed towards agents who are not distant in space but in time. The most interesting case, in my opinion, is that of blame directed towards dead persons. In my opinion, a convincing theory of blame should account for blame directed towards the dead (Coates and Tognazzini 2013). However, not everyone agrees on this point. T.M. Scanlon, for example, does not believe it necessary for his theory to account for blame towards the dead. Especially in cases of historical figures who have been dead for a long time, Scanlon believes that one cannot coherently speak of blame (Scanlon 2008, 145-146); all that is needed, in such cases, is disapproval. Blaming deceased people is thus a more controversial case than the other data points listed so far. In any case, I do not consider it to be implausible. At least intuitively, it seems that we are prone to blame at least certain agents who are not alive anymore. This is true especially if the consequences of their actions are still present and felt by others.<sup>27</sup> I concede that blaming historical figures such as Roman emperors or ancient

---

<sup>24</sup> I suspect that overt self-blame is not a widespread phenomenon. Nonetheless, our actual practices include cases of overt self-blame. Agents may feel like "kicking themselves" after realizing their faults, and sometimes they may openly manifest this feeling with words and gestures (Björnsson 2022, 199).

<sup>25</sup> For an in-depth discussion of self-blame, see Carlsson 2022.

<sup>26</sup> Shoemaker and Vargas use the term "gossipy" blame to refer to forms of what I have called "third party" blame (Shoemaker and Vargas 2021, 590).

<sup>27</sup> Just to mention an example, even Scanlon thinks that in the case of Hitler we can speak of proper blame (Scanlon 2008, 147).

tyrants may be a more controversial case. After all, these agents died so long ago that it is hard to think that we are still experiencing the negative consequences of their actions. However, I think that a convincing theory of blame should account for this phenomenon, and preferably clarify our inclination to blame recently deceased people more than historical figures from the past.

- Dispassionate blame: This is the most controversial point on this list. In fact, as I will discuss at length in the next chapter, an entire current of studies on the nature of blame does not agree to include dispassionate blame in the list of data points that a convincing theory of blame should account for. Emotional theories of blame inspired by the work of P.F. Strawson maintain that blame can be identified with an emotional mental state such as resentment, indignation, and guilt (Strawson 2008; Wallace 1998; Wallace 2011). Proponents of dispassionate blame, on the other hand, recognize the possibility of blame in the absence of any emotional mental state associated with blame. Despite the relevance of emotional accounts of blame, it seems possible to imagine cases in which the blamer does not experience the blaming emotions. As Shoemaker and Vargas notes, a mother can blame her son with nothing but exhaustion (Shoemaker and Vargas 2022, 583). At the same time, we may blame a close friend by mildly reproaching her, without experiencing emotions like anger, resentment, or indignation. And, perhaps, when we blame ourselves for our minor faults, we do not experience guilt, but annoyance. Clearly, dispassionate blame is a controversial point that not everyone would consider meriting its place among the *desiderata* of a theory on the nature of blame. However, an increasing number of philosophers started considering dispassionate blame among the data points that a convincing theory of blame should account for (Scanlon 2008; Bennett 2013; Coates and Tognazzini 2013; Fricker 2016; Sliwa 2021; Shoemaker and Vargas 2022).



In this section, I provided a list of data points that a convincing theory of blame should account for. In the second and third chapter of my dissertation, I will frequently refer to the points of this list. This is because in these chapters I will discuss both content-based theories on the nature of blame and functional accounts, comparing their ability to address the data points of the list I have provided in this section.



## CHAPTER TWO

### CONTENT-BASED THEORIES OF BLAME

As I have shown in §2, Chapter One, we can identify two possible solutions to the problem of the nature of blame. First, we have content-based accounts of the nature of blame. According to these accounts, it is possible to identify a single mental state – or a set of mental states – which constitutes blame. Second, we can refer to functional theories of blame. These theories do not identify blame with a single mental state; rather, they claim that blame is whatever mental state serves a particular function.

In this chapter, I present the most influential content-based accounts of the nature of blame. Following Coates and Tognazzini, I distinguish three main solutions adopted by proponents of content-based accounts of blame (Coates and Tognazzini 2021). In §1, I discuss cognitive accounts of blame. Proponents of cognitive accounts of blame think that blame is a cognitive mental state such as a belief or a judgment. This implies that emotions are not necessary components to the concept of blame. In §2, I take into account emotional theories of blame. According to these accounts, dispassionate blame is impossible. Emotional accounts of blame maintain that blame is, or is mostly constituted by, an emotion. Traditionally, this position has been strongly influenced by P.F. Strawson's *Freedom and Resentment*. In §3, I introduce conative accounts of blame. Conative accounts of blame try to find an answer to the problematic aspects of both cognitive and emotional theories. According to conative accounts, blame is constituted by a conative mental state such a desire or intentions. After having briefly introduced the conative approach, I discuss G. Sher's account of blame. Sher argues that blame has two components: the first element is a belief, while the second is a desire (Sher 2006). Despite the attractiveness of this position, I will argue that Sher's account is unconvincing. In §4, I present and discuss what I take to be the most plausible content-based account of blame. In his 2008, T.M. Scanlon develops a theory

of blame that is focused on human relationships. Scanlon's account of blame can correctly address most of the varieties of blame. However, even this account presents a number of problems.

Before proceeding, let me add a brief note about the methodology I follow in this chapter. This chapter does not provide an historical reconstruction of the theories of blame that have been proposed over the years. The aim here is to show and discuss from a theoretical point of view the different approaches to the analysis of blame. The overarching aim of this chapter is to show that, as long as content-based theories try to identify blame with a single mental state (or a combination of mental states), the project of finding a plausible solution to the problem of the nature of blame is doomed to fail.

### **§1 Cognitive Accounts of Blame**

Cognitive theories of blame claim that blame mainly consists in a judgment about an agent. This is not to deny that usually, when we blame someone, we feel emotions such as resentment and anger. What it means, however, is that cognitive accounts of blame tend to emphasize cognitive elements of blame, making them the core feature of our blaming practices (Coates and Tognazzini 2013, 8-10). Cognitive accounts of blame agree on the fact that blame, even when understood as a judgment on the conduct of an agent, has to involve something more than the assessment of causal responsibility. Borrowing an expression from G. Watson, judgments of the kind described by cognitive accounts of blame are judgments about "the quality of the other's moral self as exemplified in action and attitude" (Watson 2004, 226). Making use of Strawson's terminology, we can think of blaming practices as judgments about the ill will of an agent as displayed by his or her actions (Strawson 2008). To give an example, if you break a promise you made with me, my moral blame towards you would be best described as a judgment about your moral conduct. Even if it is true that I may feel anger or resentment, cognitive accounts of blame agree on the fact that these

emotions – and every other kind of emotion – are not necessary components of blame. Again, blame is best described as a judgment, or a set of judgments. Different accounts of blame identify different judgments or set of judgments in order to account for our blaming practices. Because they can differ so much from one another, in this section I will discuss two different cognitive accounts of blame. The first one is the well-known J.C.C. Smart's account of blame, which is deeply intertwined with Smart's utilitarian conception of moral responsibility. I think that, albeit not being recent, Smart's theory of blame offers me the opportunity to introduce the key concepts of any cognitive accounts of moral blame. The second cognitive account of blame that I discuss in this section is P. Hieronymi's one. Hieronymi's theory is a recent and refined cognitive account of blame that seems to find an answer to the problems faced by many other cognitive accounts. After having described and discussed Smart's and Hieronymi's accounts, I will focus on the main benefits of these theories. Before moving to the next section, I will cast doubts on the possibility of building a plausible cognitive account of blame.

A traditional cognitive account of blame is that of J.C.C. Smart. In his *Free-Will, Praise and Blame* (1961), Smart is mostly interested in the development of an account of moral responsibility rather than in the analysis of blame. Even so, the last pages of his work offer an interesting point of view on blame (and praise, too). In the mentioned paper, Smart engages with the philosophical debate about the concepts of free will and moral responsibility. Since the aim of this chapter is to offer an overview of the theories about the nature of blame, I will not spend much time discussing Smart's account of free will and moral responsibility. This does not mean that they are not relevant topics, though. On the contrary, a clear understanding of what Smart thinks about moral responsibility is necessary in order to understand his treatment of blame and praise. Smart assigns an important role to expressions such as “he could have done otherwise” or “he could not have done otherwise”

(Smart 1961, 302). In contrast to an important philosophical tradition, namely libertarianism,<sup>28</sup> Smart does not think that these expressions are incompatible with determinism. According to him, our ordinary use of expressions such as “he could have done otherwise” has nothing to do with metaphysical concerns about free will and determinism. What we have in mind when we use such expressions are ascriptions of responsibility, and in order to understand what it means to ascribe responsibility Smart asks us to imagine an example.

Suppose Tommy at school does not do his homework. If the schoolmaster thinks that this is because Tommy is really very stupid, then it is silly of him to abuse Tommy, to cane him or to threaten him. This would be sensible only if it were the case that this sort of treatment made stupid boys intelligent. [...] Now suppose that the reason why Tommy did not do his homework is that he was lazy. In such a case the schoolmaster will hold Tommy responsible, and he will say that Tommy could have done his homework (Smart 1961, 302).

Smart thinks that only in the second case Tommy is responsible for not having done his homework. If Tommy were only “really very stupid”, ascriptions of responsibility would not have any kind of effect on Tommy’s behavior. From this passage we can conclude that according to Smart ascriptions of responsibility have a pragmatic justification: “When [Tommy’s] negligence is found out, he is not made less likely to repeat it by threats, promises, or punishments. On the other hand, the lazy boy can be influenced in such ways” (Smart 1961, 302). In other words, we are justified in holding someone responsible only if he or she could be influenced by ascriptions of responsibility. With this in mind, we can start grasping Smart’s treatment of praise and blame. He begins by distinguishing two different ways in which we use the term “praise” (Smart 1961, 303). On one hand, praise could be understood as the opposite of dispraise. According to Smart, to praise or dispraise someone in this way “is simply grading a person as good or bad in some way” (Smart 1961, 303). Smart also claims

---

<sup>28</sup> See Kane 1996 for a well-known libertarian defense of free will.

that in this sense to praise or to dispraise someone is not different from grading something like apples when we claim that some apples are good, or that they taste good (Smart 1961, 303). When I say that G. Kasparov is a great chess player, I am praising Kasparov, and when I say that my best friend is a disaster in the kitchen, I am dispraising my best friend for his culinary incompetence. But the latter is different from claiming that I am blaming my friend. After all, he may have been raised in a house where he could not experiment with any recipes, or he could not have access to many different ingredients. Smart seems to understand these conditions as preventing me from blaming my friend, but not from dispraising him for his culinary incompetence (Smart 1961, 304). Just like in Tommy's example, we need to distinguish two different scenarios. If my friend could have practiced his ability in the kitchen, but he had always been lazy, then we could blame him. But if he did not have any kind of possibility to improve himself in the kitchen, then we could only dispraise him. One of the most relevant upshots of this argument is that the only difference between dispraise and blame is that blame implies responsibility: "a clear-headed man will use the word [...] 'blame' just like the previous 'dispraise', with one proviso" (Smart 1961, 304). The proviso, as we already know, is that we ascribe responsibility for his or her actions to the target of our blame. This means that according to Smart, blame consists in two distinct judgments. Precisely, blame consists in the ascription of responsibility of an agent and in a grading of that agent (Smart, 1961, 305). Smart then concludes his paper by claiming that "we should be quite as ready to *grade* a person for his moral qualities as for his non-moral qualities, but we should stop *judging* him (unless "judge" just means "grade", as in "judging apples")" (Smart 1961, 306). It is now time to highlight the elements of Smart's account of blame that are most relevant for my research. If blame consists in an ascription of responsibility and in a grading of an agent, then emotions are not a necessary component of blame. "Blame in this sense can be just as dispassionate as dispraise of a woman's nose: it is just a grading plus an ascription of responsibility" (Smart 1961, 305). This could have some advantages over

other accounts of blame that I will show in the next sections. A cognitive account of moral blame such that of Smart does not require emotions as necessary conditions for blame. This way, it can explain situations in which someone blames an agent without experiencing the so-called blaming emotions (Wallace 2011). Cases of dispassionate blame, as I have dubbed them in the previous chapter, are not so uncommon after all. So, in this respect, Smart's account of blame has an explanatory advantage on accounts of blame that identify blame with an emotional reaction to an agent's wrongdoing. In addition, Smart highlights the importance of the consequence that blame may have on the receiver. Blame, Smart claims, has the important secondary function – the primary being grading – to discourage certain behaviors. This way, Smart's account seems to explain the unpleasant aspect of blame through its function: it is so unpleasant to receive blame that those who are blamed are discouraged to act wrongly again. In addition, not only the blamed are discouraged to act in such a way. One who witnesses another agent being harshly blamed may be discouraged from acting wrongly. Thus, blame may serve an important social function, that is preventing agents from acting wrongly by discouraging some courses of action. In this framework, self-blame is explained in the same way as blame directed towards others. When I blame myself, I am grading myself while ascribing moral responsibility for the relevant action to me. Blame directed to agents distant in time and/or space, and even to dead agents, is explained in the same way: grading the agent while ascribing him or her moral responsibility. In addition, Smart's account can address the blame directed to the wrongdoer and the blame directed to a third party in the same way. In Smart's cognitive account, the person being blamed does not need to be directly addressed by the blamer: again, I can blame a third party without her being present if blame were only a grading along with an ascription of responsibility.

Smart's account certainly captures the evaluating aspect of blame, while at the same time it successfully addresses some of the data points shown in §5, Chapter One. Despite its



strengths, Smart's account has some weaknesses too. Let's start with a brief discussion on the distinction between expressed blame and unexpressed blame. As I have shown, Smart claims that blame has the secondary function of discouraging agents from acting in certain wrong ways. If one keeps blame private, however, how can blame serve this function? Of course, if I do not express blame towards anyone, it would be odd to claim that I can influence the behavior of other agents. Replying to this objection by saying that I can at least modify my own behavior is not satisfying. It seems that I can modify my future behavior even without privately blaming the wrongdoer. Thus, blame could be entirely bypassed if my intention is to modify my future conduct. Smart could reply that the primary function of blame is grading, but I suspect that his own utilitarian framework would suffer from admitting that private blame does not serve the secondary function of blame, that is influencing the behavior of agents. Unfortunately, Smart's account of blame suffers from an even stronger criticism. As we already know, he describes blame as a form of grading along with an ascription of responsibility. His conception of responsibility, however, is forward-looking:<sup>29</sup> by 'responsible agent' he means an agent whose behavior can be influenced. In turn, blame is a way of influencing the behavior of agents. But this contrasts with our blaming practices: when we blame agents for their wrongdoings, we do not primarily blame them in order to modify their future behavior. We blame agents for what they have done, and not to modify their future courses of action. Smart's forward-looking conception of responsibility and blame seems to miss an important feature of blame, namely the fact that blame is a reaction to past performances of agents. In Smart's own framework, blame seems deprived of its characteristic normative force that distinguishes it from the mere recognition that an agent acted in the wrong way. To borrow an expression from G. Watson, it apparently seems that when we blame someone (and especially morally blame someone), we act like moral

---

<sup>29</sup> According to forward-looking conceptions of responsibility, ascriptions of responsibility are justified by the beneficial consequences that can be obtained by engaging in these practices (Talbert 2019).

clerks recording moral faults (Watson 1987b, reprinted in Watson 2004, 226-227). Thus, interpreting moral blame as a form of grading plus an ascription of responsibility does not seem a good way to explain our blaming practices. Does this mean that the project of identifying blame with a cognitive mental state, such as a judgment, is doomed to fail? Not really. For now, I have only showed arguments against Smart's own account of blame, which in turn hinges on his conception of responsibility. It may be possible to build a cognitive account of blame without appealing to a forward-looking conception of responsibility. Thus, a more refined version of a cognitive account of blame could be a better explanation of our blaming practices. In what follows, I will show a promising cognitive account of blame. In the last part of this section, however, I will cast some serious doubts on the plausibility of any cognitive account of blame.

In recent years, P. Hieronymi has argued in favor of a cognitive account of blame that focuses on the normative force of blame. She expresses the exact same concern about the force of blame that I have shown before when she claims that "faced with the idea of 'mere description', one might become confused and start to worry about how a mere statement of descriptive fact can have any special force". (Hieronymi 2004, 122). So, Hieronymi too thinks that an important challenge for any cognitive account of blame is that of capturing the normative force that blame typically has. Despite this, she ultimately claims that her cognitive account of blame can account for the force of blame. According to Hieronymi, it is not necessary to look further than judgments in order to account for the normative force of blame. Hieronymi calls our attention to the importance of relationships between human rational beings. She is probably right when she claims that standing in relations of mutual regard is of considerable importance to creatures like us (Hieronymi 2004, 124). If this is right, she argues, it means that we care about what others think of us, especially if they think that we show ill will in our actions and in our behavior. But this is to say that

we care about *judgments* of other people, even if these judgments are merely descriptive. Thus, Hieronymi concludes that “the force of a judgment of ill will [...] derives from the importance of standing in relations of recognition of mutual regard. The force of a judgment of ill will is found in and carried over from its content, even if the content is merely descriptive” (Hieronymi 2004, 124). Differently from Smart, Hieronymi does not identify blame with a form of grading plus an ascription of responsibility. Even if it is true that the content of blame is a judgment regarding an agent’s ill will, Hieronymi indirectly suggests that the whole analysis of blame would include something different from a judgment – or a set of judgments. A commitment to morality and to relationship of mutual regard is necessary in order to explain the properties of moral blame, too.<sup>30</sup> While claiming that judgments alone can account for the characteristic force of blame, she notes that a commitment to morality is necessary in building a convincing account of moral blame (Hieronymi 2008, 29). What she does claim in her 2004, however, is that blame does not involve something more than a descriptive judgment, whether this additional element consists in emotions, actions, attitudes, desires, etc. Hieronymi’s account of blame surely captures the evaluational element of blame, while preserving the characteristic force of moral blame. In this respect, it has an advantage over Smart’s own approach to moral blame. Moreover, while Smart’s account of blame seems to fail in explaining the relevance of our past actions for our blaming practices, Hieronymi’s does not identify blame with an entirely forward-looking practice. Despite the advantages that Hieronymi’s account has over Smart’s, it is not entirely convincing for at least two reasons. The first reason directly addresses Hieronymi’s account of moral blame, while the second is directed to every cognitive account as well. As we have already seen, Hieronymi focuses on the normative force of moral blame, claiming that descriptive judgments alone can account for it. This leads Hieronymi to claim that emotions, attitudes and actions are not

---

<sup>30</sup> Hieronymi focuses on moral blame. As I have noted in the first chapter, my aim is to build a theory of blame that can also take into account cases of non-moral blame.

necessary to moral blame. At the same time, however, she argues that the normative force which is carried over, even by descriptive judgments alone, derives from the fact that we care about standing in relationships of recognition of mutual regard (Hieronymi 2004, 124). If this is so, then, relationships between persons seem to play a fundamental role in the analysis of blame. At this point, however, why should we keep focusing on descriptive judgments rather than directly enquiring about the role played by our interpersonal relationships? After all, here it seems that relationships are doing the heavy explanatory lift. Thus, an account of blame which focuses on interpersonal relationships seems to have better chances in understanding our blaming practices.<sup>31</sup> Even if we concede that the force of moral blame could be carried over only by descriptive judgments, Hieronymi still has to provide an explanation of how and why relationships between persons are so relevant in the analysis of blame. Without this element, her cognitive account of blame is at best incomplete. As anticipated, this is not the only problematic aspect of her account. The next problem that I will show, however, does not only affect Hieronymi's account, but it is directed towards every cognitive account of moral blame.

An example could be useful to introduce my argument. Mr. White is part of a criminal organization. This organization has many members other than White. One of these members is Mr. Blue, who is well known among criminals. White knows everything about Blue's criminal deeds, and he even admires his talent for not getting caught by the police. In the past years, White had taken several philosophy classes and courses. Among these, he really appreciated those focusing on ethics. He knows that what Blue does for his boss is wrong, and he judges Blue to be blameworthy. At the same time, White, who has an evil mind, does not blame Blue: on the contrary, he even admires Blue for his ability displayed in criminal acts. Cases like this highlight the fact that a judgment of blameworthiness is distinct

---

<sup>31</sup> In §4, I will discuss the relationship-based account of blame defended by T.M. Scanlon (Scanlon 2008).

from actual blame. But if this is so, then cognitive accounts of blame face a problem – namely, accounting for the distinction between judgments of blameworthiness and actual blame. If blame consists in judgments or beliefs, how can cognitive accounts of blame preserve the distinction between judgments of blameworthiness and actual blame? Judging an agent to be blameworthy means to recognize that blaming her would be appropriate.<sup>32</sup> But according to cognitive accounts of blame, blame does not consist in anything more than judgments – again, no emotions, attitudes, actions. Now imagine yourself thinking “my colleague is such a liar; he is not to be trusted”. Are you blaming your colleague, or merely recording some characteristics about your colleague? If blame only consisted in a judgment – or in a set of judgments – then there would be no way of distinguishing the two possibilities. It seems plausible, however, to claim that your colleague is a liar without blaming him, even if you think that he is blameworthy.<sup>33</sup> Maybe you do not care about the fact that he uses to tell lies, or you are not affected by his wrong behavior. You may even benefit from the fact that he is a liar.<sup>34</sup> Nonetheless, you judge him to be blameworthy. A cognitive account of blame, however, does not have any means to distinguish between judgments of blameworthiness and actual blame. A defender of a cognitive account, however, could say that the distinction between blameworthiness and blame still obtains. The defender of a cognitive account of blame might continue, saying that the proposition “my colleague is a liar” has two different meanings. One is a mere recording of a characteristic my colleague displays, while the other is the content of a judgment which is in turn identified with a form of moral blame. But how to distinguish the two different meanings? Remember that cognitive accounts of blame identify blame with judgments or beliefs, without the need to

---

<sup>32</sup> Note that I use the term “appropriate” or “fitting”, and not “right”. This way, I try to avoid the Moralistic Fallacy highlighted by D’Arms and Jacobson 2000. For a discussion of this fallacy applied to blame, see Graham 2014.

<sup>33</sup> Pickard argues that a judgment of blameworthiness is neither a necessary nor sufficient component of blame and of our blaming practices (Pickard 2013).

<sup>34</sup> In his 2012, D. McKenna proposes an interesting example in which someone benefits from someone else’s wrongdoings.

add any element to them – no emotions, no attitudes, no actions. So, one could not say that blaming my colleague consists in the judgment “my colleague is a liar” together with an emotion such as anger or resentment. The same goes for the judgment “my colleague is a liar” together with a modification of my attitudes towards my colleague. If this is so, however, it seems that cognitive accounts of blame do not have any instrument to distinguish between judgments of blameworthiness and actual blame. But blame seems to involve something more than a mere recording of the agent’s faults. Recall the quotation from Watson, where he claims that according to cognitive accounts of blame the complexities of our blaming practices seem to be reduced and impoverished (Watson 2004, 226-227). Again, this would suggest that, while blaming agents, we act as “moral clerks” who merely record wrongdoings of agents. So, despite Hieronymi’s efforts, cognitive accounts of moral blame seem to fail in accounting for an important characteristic of blame – namely, the fact that blame is not a mere recognition of some agent’s faults, but it is also about our own attitudes towards that agent. In other words, the blamer actively engages himself or herself while addressing the wrongdoer. If one finds cognitive accounts of blame unable to account for the personal involvement of the blamer, then he or she might think that adding emotions to judgments could be the solution. In the next section, I will further develop this idea.

## **§2 Emotional Accounts of Blame**

As we have already seen, cognitive accounts of blame tend to emphasize the centrality of judgments in our blaming practices, concluding that emotions do not play a crucial role in them. Contemporary cognitive theories of blame usually claim that it is plausible that while we blame an agent, we also feel emotions such as rage or resentment. What they insist on, however, is that emotions could accompany our blaming practices, but they do not play a central role. Thus, at least in principle, emotions are not a necessary component of blame. If one wants to provide a complete analysis of blame, emotions would not be neither necessary

nor sufficient conditions. As I have showed in the last section, however, cognitive accounts of blame seem to fail in accounting for the variety of blame. So, one could think that this failure lies in the absence of emotions. What if emotions played a crucial role in our blaming practices? In this section, I will explore this possibility, outlining the main characteristics of emotional accounts of blame. As a preliminary remark, I want to highlight a similarity between cognitive accounts of blame and emotional ones. Despite being very different, both kinds of theories assume that blame can be identified with a mental state (or a set of mental states). Cognitive accounts of blame claim that blame can be identified with a judgment (or a set of judgments). Emotional accounts, on the other hand, identify blame with an emotion (or a set of emotions). In this section, I outline what is considered to be the most influential emotional account of blame, P.F. Strawson's theory of moral responsibility and blame. A comment on the improvements R.J. Wallace has offered to Strawson's account of blame will follow. Before moving to the following section, I criticize the project of building an emotional account of blame, claiming that such an account is unable to address some relevant features of our blaming practices.

P.F. Strawson's *Freedom and Resentment* is often considered to be a landmark in studies on moral responsibility and blame (Strawson 2008). Its relevance is due to the original viewpoint from which Strawson reflects on our practices of holding agents responsible for their behavior. The philosophical tradition has often considered our practices to be dependent on facts about the agent, e.g., whether the agent acted freely or was forced to act in that way. If an agent is free to act and spontaneously decides to act in that way, then she is responsible for her action. Moral responsibility and our practices of holding others responsible, in other words, depend on facts about the agent who acted. Whether or not an agent is responsible for her action is a fact that arises from other intrinsic facts about the agent (Tognazzini 2013). Among these facts, many refer to the metaphysics of free will: could

the agent have done otherwise? Was she free to act? Was she fully determined to act in such a way by the laws of nature? Different theories of moral responsibility identify different requirements that must be met in order to be a responsible agent. What is important, though, is that they all share one feature: the existence of facts that at least partially determine whether or not an agent is responsible. If certain conditions are met, then, our practices of holding agents responsible are legitimate. Strawson's provocative idea is to reverse this conceptual order. According to Strawson, our practices of holding agents responsible determine the facts about the moral responsibility of agents. In other words, there are no facts about moral responsibility that are independent from our practices. Strawson takes our practices of holding agents responsible to be at the core of the notion of responsibility. In this way, Strawson's account of responsibility is taken to be focused on actual interpersonal practices instead of on metaphysical facts about agents. Note that until this point, I have not mentioned blame once. Why then Strawson's essay is considered to be so important for studies on blame?

The answer lies in the way Strawson treats interpersonal practices of holding agents responsible. According to Strawson, it is a fundamental fact about the human condition that we react to agents who display ill will towards others. Among these reactions we can find emotions such as resentment or anger. These reactions are the reactive attitudes, which are at the core of interpersonal practices. Strawson himself does not provide a full list of the reactive attitudes: gratitude, resentment, love, hurt feelings are all labeled as reactive attitudes (Strawson 2008, 5). Importantly, reactive attitudes are not limited to personal reactions towards agents who wronged us directly. Strawson claims that certain reactive attitudes are essentially vicarious or impersonal: these reactive attitudes are "reactions to the qualities of others' wills, not towards ourselves, but towards others" (Strawson 2008, 15). Finally, Strawson also mentions the possibility of directing the reactive attitudes towards ourselves,



as in the case of guilt. The analysis of our actual practices of holding agents responsible leads us to the reactive attitudes, which are emotional responses to agents' display of good or ill will.<sup>35</sup> So, in the Strawsonian account that many seem to accept, being a responsible agent means being an apt target of reactive attitudes. We now have a definition of moral responsibility, but what about blame? Surprisingly, Strawson himself only mentions blame six times throughout his paper. Despite this, philosophers working on blame consider the identification of the reactive attitudes with blame to be much plausible.<sup>36</sup> It is thus safe to say that according to Strawson – or at least according to the vast majority of his interpreters – to blame an agent consists in targeting that agent with the reactive attitudes. Note that interpreting the reactive attitudes in this way leaves us open the possibility of defining blameworthiness, too. According to this Strawsonian take on moral responsibility and blame, a blameworthy agent is an appropriate target for the reactive attitudes. What are the advantages of the Strawsonian account of blame? First, recall the criticism directed against cognitive accounts of blame. As we have already seen, cognitive accounts tend to blur the distinction between blameworthiness and blame, a distinction that we may want to maintain. Emotional accounts of blame such as that of Strawson, on the contrary, do not face the same problem. According to the Strawsonian account of blame, an agent is blameworthy when she is an apt target for the reactive attitudes. Active blame, on the other hand, requires that the blamer actually feels an appropriate emotion such as resentment or indignation towards the blamed. This is not the only advantage of an emotional account of blame. Emotions could be kept private or be expressed, thus emotional accounts of blame can address both the phenomena of private and overt blame. As we have seen, moreover, overt blame could

---

<sup>35</sup> It could be argued that the identification of the reactive attitudes with emotional responses is not obvious. Strawson himself, however, seems to interpret reactive attitudes as emotional responses. For reference, see the conclusion of his essay, where he claims that: "It is a pity that talk of the moral sentiments has fallen out of favor. The phrase would be quite a good name for that network of human attitudes" (Strawson 2008, 25).

<sup>36</sup> In his 2008, Scanlon claims that this is a "natural interpretation of the reactive attitudes; an interpretation that many seem to accept and that I do not want to argue against" (Scanlon 2008, 224, footnote 6).

be directed towards the wrongdoer or it may involve third parties, and, again, emotional theories of blame such as that of Strawson do not have difficulties in addressing this distinction. We could express reactive emotions against a wrongdoer who directly offended us, or we could express them against someone who did not mistreat us directly. In addition, emotional accounts of blame have the important merit of acknowledging the backward-looking aspect of blame: when we blame someone, we are reacting to the wrongdoer with emotions such as resentment or anger. We are not blaming him or her in order to control their future behavior, but we are blaming the wrongdoer for what he or she has done. It seems then that emotional accounts of blame could address the variances of our blaming practices: does this mean that we have finally found a convincing theory on the nature of blame? Further considerations will show that it is premature to claim this.

Claiming that blame consists in targeting an agent with a reactive attitude (or a set of attitudes) implies that there cannot be instances of what I dubbed before as ‘dispassionate’ blame.<sup>37</sup> With this term, I refer to blaming practices which do not involve any kinds of emotional responses directed towards the blamed. According to emotional accounts of blame, no blaming activity takes place without an appropriate emotion, e.g., resentment. Despite its initial plausibility, this last claim is unlikely to be true. As I have noted in Chapter One, dispassionate blame seems to be among the data points that a theory of blame should account for. Moreover, think of situations in which we blame historical figures from the past. When studying the history of the Roman empire, we may not feel anything at all towards figures such as Heliogabalus or Nero. We may not feel any emotions towards them precisely because they are so distant in the past. From this, however, it does not follow that we cannot blame them. I could think that Heliogabalus and Nero acted in horrible ways and blame them for their cruel treatment of citizens of the Empire. More generally, our practices of blaming

---

<sup>37</sup> See §5, Chapter One.

agents distant in time or space seem hard to explain within an emotional account of blame. Naturally, a defender of an emotional account could reply arguing that in such situations we are not blaming those agents, but only judging them to be apt targets of reactive attitudes. But this may seem an *ad hoc* reply. This point is made even clearer by G. Sher, who enumerates situations in which dispassionate blame is at work:

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. As [these] examples suggest, blaming is something that we can do regretfully or dispassionately (Sher 2008, 88-89).

Moreover, it is important to note that not every emotion can count as blame. Emotional accounts of blame must be precise in pointing out the relevant emotions which count as blame. This should be evident when we consider people who stand in a close relationship with each other. We may feel sad or embittered because of an action performed by our loved ones, but, at the same time, we may not morally blame them. This suggests that certain emotions such as sadness do not count as blame.<sup>38</sup> A defender of an emotional account, thus, must be careful in specifying which emotional states count as the reactive attitudes. Even when philosophers agree on the importance of emotions in accounting for blame, they tend to have different ideas about which kind of emotion(s) an emotional account of blame should consider.<sup>39</sup> R.J. Wallace, whose reading of Strawson is one of the most accurate and interesting, claims that only resentment, indignation, and guilt count as reactive attitudes relevant for blame (Wallace 1998, chapter 2). According to Wallace, Strawson is not precise

---

<sup>38</sup> D. Pereboom would strongly disagree with this sentence, for he has built an account of blame which replaces emotions such as anger and resentment with sadness (Pereboom 2013). His account of blame is deeply connected to his theory of free will and moral responsibility, which are not the focus of my research. Thus, I do not discuss his account here. For my purpose, it is sufficient to note that despite his controversial claim that sadness is the appropriate emotion linked to blame, we usually think of blame as associated with anger, resentment, indignation or similar emotions.

<sup>39</sup> D. Shoemaker, for examples, considers anger to be fundamental to our blaming practices (Shoemaker 2015; 2017).

enough in characterizing which attitude counts as a reactive attitude. But this is an important point, since, as we have seen, blame could not consist in every kind of emotion or attitude. This leads me to the discussion of a potential problem of emotional accounts of blame.

Strawson claims that the reactive attitudes include the full range of feelings we are susceptible to in interpersonal interactions. But this interpretation of the reactive attitudes risks being overinclusive. As Wallace rightly points out, “the wider we stretch the class of reactive attitudes making them coextensive with the emotions we feel toward people with whom we participate in interpersonal relations, the less plausible the claim becomes that holding people responsible is inextricably a part of this web of attitudes” (Wallace 1996, 11). This is true not only for holding agents responsible, but also for blaming them. Claiming that the class of reactive attitudes includes the full range of emotional responses to agents’ actions misunderstands an important characteristic of blame, namely its plausible connection with a belief that the agent who is being blamed acted in a wrong way or failed to meet a relevant standard. Without this kind of propositional content, an emotional account of blame would take a strongly non-cognitivist turn. Every emotional response would count as blame only because it is felt and expressed by the blamer. But this contrasts with our intuitions about blame, and it contrasts with our blaming practices too. Even if the reactive attitudes are emotional responses, they must be at least partially rationally grounded. This point is made clearer by referring to excuses and exemptions. If an agent acted in a morally wrong way, but we know that she has a valid excuse, we are not prone to blame her as though she does not have an excuse. Thus, it seems that even an emotional account of blame needs to inform blame with a propositional content.<sup>40</sup> If this is so, however, emotional accounts of blame lose part of their appeal. According to this kind of accounts, blame consists in an emotional

---

<sup>40</sup> Recent emotional accounts of blame have been incorporating this kind of propositional content (Graham 2014). For a discussion, see Macnamara 2013. Still, it seems to me that, in this way, emotional accounts of blame partly lose their initial appeal.

reaction to wrongdoers. As we have seen, however, even emotional accounts of blame admit that blame also has a cognitive component. But then, what are the specific reasons to defend an emotional account? It may be true that blame is often accompanied by hostile emotional state, but this does not prove that emotions are necessary to blame, nor that blame consists in emotional responses to wrongdoers. If emotional accounts of blame require a cognitive element, why not consider this cognitive element, instead of emotional reactions, as central to our blaming practices? If so, it becomes less clear why we should think that emotions should be privileged in our understanding of blame.

Upon reflection, it seems that interpreting blame as an emotional reaction cannot provide a full analysis of blame. If so, however, we find ourselves in an awkward position. How should we clarify the nature of moral blame, if both cognitive and emotional approaches seem to fail in their task of explaining moral blame?

### **§3 Sher's Account of Blame as a Disposition Around a Belief-Desire Pair**

Content-based accounts of blame do not need to be reduced to cognitive and emotional accounts only. As I have shown in the last two sections, identifying the core feature of blame with a descriptive judgment or with an emotion seems to fail in addressing some of the data points of blame. Thus, recently, attempts have been made in other directions. In this section and in the following one, I will discuss two different accounts of blame that are neither purely cognitive, nor purely emotional. Despite their differences, both accounts can be classified as conative accounts of blame (Coates and Tognazzini, 2013, 10). Conative accounts of blame claim that blame consists in something more than a judgment, but they do not require that the blamer feels a specific emotion (or any emotion at all). What conative accounts of blame add to judgments are elements such as desires, volitions, intentions, and generally all those elements which require an active engagement on the part of the blamer. Conative accounts of blame differ from each other according to the conative element they tend to emphasize.

In this section, I will discuss G. Sher's account of blame, while in the following one I will focus on T.M. Scanlon's.

In his 2006, Sher argues that both cognitive and emotional accounts are inadequate in understanding blame and our blaming practices. In order to understand blame (and, in particular, moral blame), we must ask ourselves why we feel and why we act as we do towards those we blame (Sher 2006, 93). But how exactly do we feel and act towards those we blame for something? As we have seen, a wide array of emotions and dispositions is usually associated with blame, at least at a psychological level. On the one hand, blame is often associated with feelings such as anger and resentment, hostile behavior, or even self-reproach. On the other hand, blame does not identify with any of the elements of this list. Thus, according to Sher, a convincing theory of blame must account for this variety of dispositions associated with blame, together with an element capable of explaining such dispositions. Drawing from cognitive accounts of blame, Sher claims that judgments or beliefs play a crucial role in explaining blame. More precisely, he thinks that the belief component of blame must be the belief that an agent has performed a wrong action for which he or she is blameworthy (Sher 2006, 9). At the same time, he thinks that judgments or beliefs alone are insufficient. Sher's strategy, then, is to find a single element – or, better, a single type of element – that, added to a judgment of blameworthiness, could explain the different blame-oriented dispositions in a unifying way. The justification of this methodology comes from Sher's belief that “in general, the preferred strategy for explaining why someone has a number of seemingly disparate dispositions is to take him to have a single goal whose achievement requires that he manifests each disposition” (Sher 2006, 99). This goal, as Sher claims, can be expressed in the form of a desire. So, for example, if I have the desire to compete in and win a weight-lifting contest, this desire will explain different attitudes and dispositions that I could manifest. Eating many highly caloric meals, training with heavy

weights and doing a lot of stretching are all attitudes that could be explained by my desire to win the competition. Sher wants to adopt the same methodology with moral blame: “if we can trace each blame-related disposition to a single desire, then we will be able to explain what those dispositions all have in common” (Sher 2006, 100-101). Thus, according to Sher, a desire must be added to the judgment of blameworthiness. But then we should ask an obvious question: which desire can explain, together with a judgment of blameworthiness, all the various blame-related dispositions? Before answering this question, let me introduce the terminology used by Sher. He calls B the belief component of his theory, namely a judgment of blameworthiness, and D the desire that should explain the blame-related dispositions. So, the best explanation of the blame-oriented dispositions should be the combination of a belief and of a desire, or, as Sher calls it, the couple belief-desire (Sher 2006, 103). Consequently, Sher asks whether “is there in fact a desire that, in conjunction with B, would give rise to each blame-related reaction under roughly the conditions in which we would in fact have it” (Sher 2006, 101). Now, let me discuss a serious potential problem in the analogy between the desires that a person may hold and the kind of desire that should explain blame-related attitudes. In the weight-lifting competition example, competing and winning are goals, which are future-oriented. But the kind of desire that should explain blame is not like goals, for it is not forward-looking. My beliefs that I should eat highly caloric meals, train with heavy weights and do a lot of stretching are all capable of meshing with my future-oriented desire to win the competition. But moral blame seems to be very different. What I believe when I morally blame someone is that he or she acted (at the past tense) in a wrong way for which they are blameworthy. If this is so, however, the kind of desire that we are looking for must be different from future-oriented goals. As Sher claims, “Because [...] B is oriented exclusively to the past or present, B does not seem capable of meshing with any future-oriented desire in a way that could account for our blame-related dispositions” (Sher 2006, 101). Thus, if we want to understand what B and D are, we must first align them

on a temporal axis. B and D must be either future-oriented, or past-oriented. Sher firstly considers the possibility that B and D were both future-oriented, but he dismisses it quickly. If we “view each blame-related disposition as traceable to the combination of a future-oriented expectation of bad conduct and an equally future-oriented desire to forestall such conduct” (Sher 2006, 102), then moral blame would be reduced to an instrument of social control. But this is not what we have in mind when we morally blame the wrongdoer. Moreover, if blame were entirely future-oriented, it would not make any sense to blame agents who do not pose any threat in the future. When we blame the dead, however, we are doing precisely so. Thus, again, taking B and D to be both future-oriented does not seem the best way to explain blame. If so, then we can think that B and D should be both oriented to the past. As I have already claimed, the belief associated with blame is that an agent acted in the wrong way. So, we can see that B is oriented to the past. Accordingly, Sher argues that what explains our blaming dispositions is the “backward-looking desire that the person *not have done* what he in fact did” (Sher 2006, 102). Now we finally have an answer to the question raised before. What is D, the desire component of blame? It is the desire that the wrongdoer has not performed the wrong action. So, now we can finally understand why Sher refers to his own account as “two-tiered” (Sher 2006, 14-15; 112-115): the belief-desire pair forms the first tier, while the blame-related dispositions form the second. To sum up, according to Sher, blame is a set of dispositions organized around a central desire-belief pair. Sher also argues that asking whether the link between the blame-related dispositions and the belief-desire pair is necessary or contingent is not relevant. After all, he wants to explain the *phenomenon* of blame as we experience it, not giving a conceptual analysis or a definition of the word ‘blame’. Thus, no need to prove that the link between dispositions and belief-desire pair is necessary. It would be pointless to refer to imaginary counterexamples where one possesses the relevant belief-desire pair without the dispositions related to blame (Sher 2006, 112-113). The reason why this move would be pointless is that Sher is interested in explaining



blame in the actual world, where every actual person who has a belief-desire pair of the relevant sort also has most – if not all – of the relevant dispositions associated with blame. According to Sher, his own theory of blame does not need to give the full analysis of blame: as long as this theory explains our blaming practices, a contingent link between dispositions and the relevant belief-desire pair is sufficient.

Sher's theory of blame does a good work in accounting for many different aspects of the phenomenon of blaming – again, the data points that I have identified in the last section of the previous chapter. The dispositions around the relevant belief-desire pair can be expressed or kept private, thus accounting for both overt and unexpressed moral blame. Moreover, the belief-desire pair could explain and justify dispositions that an agent directs towards him or herself. At the same time, it seems that Sher's conative account of blame does not identify blame with a forward-looking practice of social control, thus effectively capturing the backward-looking aspect of blame. Finally, this account does not reduce blame to a mere recognition that the blamed acted in a wrong way. The distinction between blame and blameworthiness is thus preserved, while accounting for the unpleasant aspects of blame. So, it really seems that Sher's theory has several advantages over cognitive and emotional accounts of blame. Have we finally found a convincing account of the nature of moral blame? As I will show in the following paragraph, things are not that simple.

Sher claims that blame consists in a set of dispositions organized around and explained by a belief-desire pair. It seems safe to say, then, that the belief-desire pair has an important explanatory priority in Sher's two-tiered account of blame. The dispositions which are frequently associated with blame are unified and explained by the belief-desire pair (B-D pair). As we know, B consists in the judgment that the agent is blameworthy. This kind of judgment alone is insufficient to blame, thus Sher introduces the desire that the blameworthy agent had not done what he has done. Upon reflection, we can notice that the belief that the

agent has performed a wrong act for which he or she is blameworthy is an element of blame that Sher's account shares with most accounts of moral blame (see, for example, Scanlon 2008). So, what is truly original in Sher's theory of blame is the desire component (D). D, moreover, seems to be the element of the belief-desire pair that really does the work of explaining the dispositions associated with blame. In her 2013, A. Smith raises a serious objection to Sher's account of moral blame: is it true that adding to a judgment of blameworthiness the desire that the blameworthy agent acted in a different way can capture and make appropriate all and only those reactions commonly associated with blame? Some examples proposed by Smith help understand the negative answer to this question. Think of reactions towards a loved one, or, in general, towards people with whom we entertain a meaningful relationship. You may believe that a loved one is truly blameworthy for an action he or she committed, and you may have the desire that that agent has not committed that action. So, the belief-desire pair that Sher describes is present. But it is perfectly conceivable that you may not blame that person, considering your meaningful relationship with him or her. Now, recall that Sher is not interested in giving an analysis of the concept of moral blame. His research focuses on the phenomenon of moral blame; thus, he is not concerned about imaginary counterexamples that invoke impossible scenarios. But the situation I have just described is far from being distant from our actual practices. Therefore, the belief-desire pair is not sufficient to unify and explain the dispositions frequently associated with blame. These dispositions, as we have just seen, could be absent among persons standing in certain kinds of relationship. This is not to say that in situations like these, when the belief-desire pair is present, no dispositions associated with blame could be explained by the B-D pair. We may feel sad or let down because of the actions of the wrongdoer, and these dispositions would be explained by referring to the belief-desire pair. But these are not dispositions frequently associated with blame, like anger or reproach are. So, it seems that the belief-desire pair is not sufficient to blame. But is it even a necessary component of blame? I think it is

not. Smith refers to a concrete event in order to show that the belief-desire pair – and, more precisely – the desire component is not necessary to explain our blaming practices:

Consider, for example, the attitudes of many Republicans on learning of Bill Clinton’s ill-fated dalliance with Monica Lewinsky. I think it fair to say that a 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 (Smith 2013, 35).

In fact, situations like this reveal a psychological trait which is quite common among human beings. Under certain conditions, it is not true that a desire such as that described by Sher – together with a belief of blameworthiness – explains the dispositions associated with moral blame. We sometimes blame persons for what they have done even if we do not have the desire that they have acted in a different way. As Smith argues, we may even relish the missteps of others and yet blame them for their wrongdoings (Smith 2013, 35). Thus, the desire that the agent had not acted in the way he or she actually acted does not always effectively explain the blaming dispositions. Sher is careful in stating that his theory of blame wants to explain the phenomenon of blame as it unfolds in the actual world. In the actual world, however, the belief-desire pair – which Sher considers to be central in the explanation of the blaming dispositions – is neither sufficient nor necessary to explain our blaming activities. Despite its advantages, Sher’s account of blame does not seem so attractive anymore. This is not to say that every conative account of blame would fail in explaining blame. Maybe, there is one way in which an alternative conative theory could avoid the pitfalls of both cognitive and emotional accounts, while avoiding relying on the belief-desire pair described by Sher. In the next section, I will conclude my analysis of content-based theories on the nature of blame by discussing Scanlon’s relationship-based theory of blame. Will

Scanlon's account of blame finally succeed in answering our questions about the nature of blame?

#### §4 Scanlon's Relationship-based Account of Blame

In this section, I want to focus on Scanlon's account of blame. In his 2008, Scanlon defends what we may classify as a relationship-based account of blame (Scanlon 2008, 122 onwards). While cognitive approaches to blame put beliefs at the core of our blaming practices, and emotive accounts of blame put emotions in the same place, Scanlon emphasizes the centrality of relations between people. Scanlon offers an interpretation of the phenomenon of blame which does not identify blame with a mere evaluation – as if blaming a person were equivalent to assigning 'grades' to her – nor with a form of sanction, as if blame were a mild form of punishment. According to Scanlon, the complexity of the phenomenon of blame could be understood if we interpret blame as a form of reaction towards individuals who have impaired, through their actions, the relationship in which we are with them. More precision is needed here, though, for Scanlon introduces blame by defining blameworthiness first. He writes that “a person is blameworthy for an action [when] his action indicates something about that agent's attitudes toward others that impairs his relations with them. To blame someone is to hold attitudes toward him that differ, in ways that reflect this impairment, from the attitudes required by the relationship one would otherwise have with the person” (Scanlon 2008, 145). This way of understanding blameworthiness and blame reflects the importance that Strawson's *Freedom and Resentment* had for discussions on moral responsibility and blame: Scanlon himself admits that his own account owes to Strawson the importance given to interpersonal relations (Scanlon 2008, 128). At the same time, Scanlon's account of blame differs from Strawson's because it focuses not on emotions and sentiments such as resentment and indignation, but on relationships between persons. Before giving the

details of Scanlon's proposal as well as moving objections to that, I will briefly enumerate the many advantages of his relationship-based account of blame.

Seeing blame as a reaction to behaviors that impair a relation between persons has the advantage that blame is not reduced to mere grading, and it is not identified with a form of sanction. At the same time, both private and overt blame could be explained with this account: one does not need to overtly react to an impairment of the relation between persons in order to modify oneself attitudes towards the blamed. Moreover, Scanlon can elegantly explain situations of what we may call dispassionate blame. We do not need to feel resentment or indignation in order to modify our attitudes towards the blamed in the way that this modification reflects our judgment that the blamed has acted in a way that impairs our relationship with him or her. An important feature of this account, moreover, is that it captures our intuitions about forward-looking and backward-looking conceptions of blame. On one hand, blame is a reaction to an action done in the past; on the other hand, blame is interpreted as a modification of an ongoing relationship which will have some consequences in the future. The main problems of a relationship-based account such as that of Scanlon are situations in which one blames him or herself and situations in which we blame persons who do not have, nor have had, any causal interactions with us. In the last part of this section, I will develop these critiques extensively, but let me now offer a more detailed account of Scanlon's argument.

Scanlon begins his treatment of blame by considering the relationship of friendship. This is the example that he uses: "Suppose I learn that at a party last week some acquaintances were talking about me and making some cruel jokes at my expense. I further learn that my close friend Joe was at the party, and that rather than coming to my defense or adopting a stony silence, he was laughing heartily and even contributed a few barbs, revealing some embarrassing facts about me that I had told him in confidence" (Scanlon 2008, 129). Joe's

actions reveal his ill will towards me, and they modify the status of my friendship with him. One of the advantages of Scanlon's account is that of classifying as blame many different possible reactions I may have towards Joe: from directly addressing him, to tacitly deciding not to talk to him anymore. These reactions count as moral blame, even if they are different from each other. In order to better understand his account of blame, some further details about the notion of relationship between persons are needed. Scanlon claims that a relationship is constituted by certain attitudes and dispositions (Scanlon 2008, 131). Among these, a central role is played by intentions and expectations about how the persons involved in the relationship will behave toward one another, as well as about emotions that these persons will feel towards one another. In the case of friendship, these attitudes and dispositions are also mutual: I cannot say that I am your friend, if I know that you hold certain attitudes towards me just because you know that in doing so you will get some benefits. Scanlon continues his analysis of relationships writing that "it is important to distinguish, here, between the normative ideal of a relationship of a certain kind, such as friendship, and particular relationships of that kind, which hold between particular individuals" (Scanlon 2008, 133). The normative ideal of a particular kind of relationship specifies the conditions that need to be met by a relationship in order to count as a relationship of this kind. In the case of friendship, the normative ideal of friendship specifies the conditions under which is true that a particular relationship counts as friendship. If these conditions are met, then a relationship counts as friendship. Scanlon also thinks that what he calls the normative ideal of a relationship specifies how individuals in this relationship should behave towards each other (Scanlon 2008, 134). This distinction will play a crucial role in my argument against his account of blame, but for now let us proceed by showing Scanlon's position. Scanlon claims that relationships can be impaired in different ways, some of which would justify blame from one party towards the other. I could simply move far away from where I grew up and lose some of my old friends: this is not the kind of

impairment of a relationship that Scanlon has in mind. He has good reasons to deny this possibility, for in moving far away from home I have not shown ill will towards my old friends. In this situation, blaming me would be unjustified. Scanlon claims that impairments that justify blame occur “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 involves” (Scanlon 2008, 135). In other words, when one party impairs the relationship in which he or she stands with the other party, the latter is justified to react to this impairment by holding attitudes towards the former that are normally ruled out by the nature of the relationship. This reaction to the impairment of a relationship is what Scanlon identifies as blame.

Now that I have detailed a little bit more Scanlon’s account of blame, I will highlight some critical remarks. Let us start with the problem of self-blaming. As we already know, the account of blame I am describing here is relationship-based. This means that if it is impossible to identify any kind of relationship between the blamer and the blamed, then there is no conceptual space for blame. But when I blame myself for my wrong behavior, the kind of relationship which gets impaired by my own behavior is difficult to identify. Scanlon addresses this problem by claiming that the difficulty of understanding self-blame in his relationship-based account of blame is only apparent. After all, he says, we blame ourselves for our lack of proper concern for others: “because of one’s own attitudes toward and treatment of others, one can no longer endorse one’s own feelings and actions, but must instead endorse the criticisms and accusations made against oneself by others” (Scanlon 2008, 154). To put it differently, in order to fit self-blame into his account of blame, Scanlon has to accept a kind of self-estrangement. This solution, however, could be seen as ad hoc and not very convincing: Scanlon is compelled to engage in what Shoemaker and Vargas call

“fancy dancing” (Shoemaker and Vargas 2021, 582-585) in order to address the phenomenon of self-blame. One could bite the bullet on self-blame if it helps defend a relationship-based account of blame. I suspect, however, that the following arguments will cast more doubts on Scanlon’s account of moral blame.

Wallace provides two solid arguments against Scanlon’s account of blame. Recall that according to Scanlon emotions and feelings are not needed in order to blame someone. Even if I agree with him on this, I concede that Wallace has a point when he argues that Scanlon’s account “leaves the blame out of blame” (Wallace 2011, 349). Wallace claims that blame has an element of opprobrium, and in his account of blame, emotions and sentiments play the role of being the elements of opprobrium which blame carries with itself. According to Wallace, blame - as Scanlon describes it - loses its force, its unpleasant characteristic. Scanlon can defend himself by claiming that the element of opprobrium of blame consists in the modification of one’s attitude towards the blamed in a way that is appropriate with the corresponding judgment of blameworthiness. According to Wallace, however, this is insufficient to capture the negative and unpleasant aspect of blame. After all, we can imagine a situation in which you are simply sad and frustrated by the fact that one of your friends has not kept his promise to you, and as a reaction you modify your attitudes towards that person. Now, if you modify your attitudes in accordance with these sentiments, then it would be odd to say that you are blaming that person. After all, you are only feeling sad, an emotion which is not commonly linked to blame. Wallace makes this point when he writes that “it is one thing to modify your standing intention to confide in a person who has betrayed you out of a sorrowful recognition that this attitude is no longer appropriate to the nature of the relationship between you. It is quite another to modify the attitude when you also feel warranted resentment of the person for letting you down in this way” (Wallace 2011, 358). According to Wallace, only the latter is an instance of blame, while the former would be a



modification of one's attitudes towards another person without any moral blame involved. We do not need to endorse an emotional theory of blame in order to see that Wallace may be correct in pointing out a false negative in Scanlon's account of blame. After all, it is true that I can react to an impairment of my relationship with someone while feeling sad and disappointed. This would be described as moral blame by Scanlon's account, but Wallace is right in claiming that in this situation there is no moral blame involved. Someone could think that this argument rests on Wallace's intuitions about the correctness of emotional accounts of blame. Because of this possibility, in the last part of this section I will present a stronger argument against Scanlon's position.

As we already know, Scanlon's account of blame is relationship-based. In order to defend his thesis, he discusses a common kind of personal relationship, namely friendship. When he sums up his results, he claims that five elements are central to his account of blame (Scanlon 2008, 138): there must be a ground relationship; that relationship must be impaired by one of the parties involved; the blamer must be in a certain relevant position relatively to the blamed; the impairment must have some significance for the blamer; and the blame must be appropriate. I will focus on the first two elements of this list in order to build the argument against Scanlon's account of blame. To begin with, think of a quite common situation. Most of us engage in blaming practices with persons who do not have any kind of relationship with us. But these instances of blame, the argument goes on, cannot be explained by Scanlon's account, since there is no ground relationship which could get impaired. And if there is no impairment of the relevant relationship between the blamer and the blamed, there cannot be any blame, given that blame involves such impairment. I will frame this argument as the problem of blaming the stranger (Sher 2013). Scanlon is aware of this problem, and he spends an entire section of his book dealing with it (Scanlon 2008, 139-152). His solution, however, is not convincing. Scanlon claims that the problem of blaming the stranger finds

its solution in noticing that it is false that we do not have a certain kind of relationship even with strangers. Of course, this relationship is very different from friendship. But it is a relation nonetheless, and of an important kind: it is the moral relationship. So, for Scanlon, there is a universal relationship that holds between everyone simply in virtue of us being fellow rational agents. Recall that in the case of friendship Scanlon identifies a normative ideal of friendship that specifies attitudes and expectations that we should have towards each other if we are friends. In this respect, the moral relationship is just like friendship. We can describe morality as a normative ideal which sets the standards of attitudes and expectations which we should adopt towards each other. Friendship requires that friends adopt certain attitudes towards each other because otherwise the persons involved in that relationship would not be friends at all, while morality “requires that we hold certain attitudes toward one another simply in virtue of the fact that we stand in the relation of ‘fellow rational beings’” (Scanlon 2008, 140). If so, then, Scanlon seems right in claiming that the idea of a universal moral relationship is not so implausible as it may seem. After all, the moral relationship shares with friendship and other personal relationships the common characteristic of being a normative ideal which sets the standards of attitudes and expectations that the parties involved should have towards each other. If Scanlon is right about this, then morally blaming the stranger would not be a problem anymore for his account of blame. After all, even between strangers a certain kind of relationship holds, namely the moral relationship. And if there is such a relationship, then it could get impaired by other persons’ wrong behavior, and so our reaction to such an impairment would be justified. But reactions to such an impairment are what Scanlon identifies as moral blame, and thus moral blame would be justified. As Wallace points out, however, the analogy between friendship and the moral relationship breaks down at a crucial point (Wallace 2011, 360). In the case of friendship, as I have already shown, the normative ideal of friendship identifies the conditions that must be met for a particular relationship to count as friendship (Scanlon 2008, 133–34). These

conditions, in other words, are friendship-constituting. This set of conditions and norms, however, must be carefully distinguished from another set of norms, which we can consider as friendship-based norms. Friendship-based norms specify intentions and expectations that persons who already stand in the relationship of friendship should have towards each other. Norms that are based on a relationship such as friendship are not the same kind of norms which identify the conditions that must be met for that relationship to count as a case of friendship. As Wallace states, “friendship-based norms will apply to a person only if it is true of them that they already stand in a relationship of friendship to a given person; to determine whether this is the case, we must appeal to a different set of standards” (Wallace 2011, 358). Recall that a central feature of Scanlon’s account of blame is the notion of impairment of a relationship. When a friend of mine betrays me or breaks a promise he made with me, his actions have a certain effect on our friendship. His actions alter the meaning of our friendship by going against those friendship-constituting norms that discriminate between relationships that count as friendship and relationships that do not count as friendship. My putative friend’s actions, in other words, impair our friendship by reference to friendship-constituting norms. Persons who stand in a relationship of friendship share intentions and expectations in virtue of the existence of friendship-based norms which have normative significance for friends. If your putative friend betrayed you, his action altered the meaning of your relationship, thus giving you reason to modify your attitudes even in contrast with those friendship-based norms which regulate expectations and intentions between friends. This modification of your attitudes towards your putative friend, thus, should be tracked down to the impairment of your relationship with him. The process has two steps, and it involves two different sets of norms: a friend of yours impairs the relationship altering it, and you have reasons to modify your attitudes towards him. The impairment of your relationship with him involves the norms that constitute your relationship, and this in turn will undermine the reasons and obligations that are based on the nature of that relationship. As I will show now,

the same model cannot be applied to the moral relationship. Scanlon thinks that the universal moral relationship holds between every rational human being simply in virtue of the fact of sharing a property, namely the property of being rational agents. So, what constitutes the moral relationship is our sharing the property of being rational agents. All rational agents stand in the moral relationship, which in turn grounds a series of expectations and intentions. It seems that we can find the same structure that we find in friendship: someone – even a stranger – alters the relationship-constituting norms and conditions, and, in response, I modify my intentions and expectations based on the moral relationship. This analogy, however, is wrong. For no matter what kind of morally wrong action is performed by someone who does not have any kind of relationship with me other than the moral relationship, that action could not impair the relationship-constituting norms (Wallace 2011, 362). The stranger who treats me with disregard cannot alter the fact that we two remain rational agents. But if being rational agents is what constitutes the moral relationship, then the moral relationship cannot be impaired. If this is so, however, Scanlon’s account of moral blame cannot justify blaming the stranger by invoking the moral relationship. In other words, given Scanlon’s treatment of the moral relationship, no impairment of this relationship is possible. But blame is a reaction to an impairment of an ongoing relationship between persons. If there is no conceptual space for impairment, then blame cannot be justified. In conclusion, Scanlon’s account of blame fails to address the problem of blaming the stranger, and thus fails to account for an important characteristic of moral blame, namely the possibility of morally blaming someone with whom we do not stand in any kind of personal relationship. Scanlon could defend his theory by claiming that when a person who stands in the moral relationship with you treats you with disregard, there still is something that gets impaired. What gets impaired here is the capacity of the wrongdoer to relate to you in a way, which is characteristic of mutual recognition (Wallace 2011, 363-364). Scanlon himself, however, admits that there is nothing one could do – no matter how wrong it is – which will

justify a revision of your attitudes towards the wrongdoer such that it will involve the violation of basic moral rights (Scanlon 2008, 142). But this element strongly suggests that the moral relationship is insufficient, and that we need to refer to something external to the moral relationship – namely, to morality itself and its normativity. If this is so, however, moral reasons are not based on the relationship we already stand in, but they need to refer to something external to the relationship. If so, then, why should we suppose that a universal moral relationship holds in the first place? Scanlon proposed the idea of the existence of a universal moral relationship as a solution to the problem of blaming the stranger in an account of moral blame that is relationship-based, but I hope I have showed that the moral relationship fails to do so.

In the end, it seems that even Scanlon's account of blame cannot address the variety of blame that I have highlighted in Chapter One. It seems that cognitive theories of blame, emotional theories of blame, and conative theories cannot give a convincing explanation of blame and of our blaming practices. All these accounts of blame have in common the identification of blame with a mental state, or a set of mental states. Thus, maybe, this is the problematic aspect. In other words, content-based theories on the nature of blame are unable to account for the complexities of our blaming interactions. In the next chapter of this dissertation, I focus on a different approach to the problem of the nature of blame. What if we adopt a functional point of view, and try to define blame through its function?

## CHAPTER THREE

### FUNCTIONAL THEORIES OF BLAME

In this chapter, I present and discuss an alternative approach to the study of blame. According to such an approach, the problem of the nature of blame can be solved by analyzing the function of blame. As I have noted in the previous chapter, content-based accounts consider blame as a mental state or a set of mental states. Thus, blaming reactions can be distinguished from other kinds of reactions by identifying the occurrence of a mental state or a set of mental states. Functional theories of blame, conversely, focus on the function discharged by our blaming practices – it is the function of blame that determines which attitudes and activities count as instances of blame (Shoemaker and Vargas 2021, 587). Functional accounts hold that whichever mental state or activity that discharges a particular function count as blame (Coates and Tognazzini 2021). Emotions such as resentment and indignation, beliefs, attitudes, desires, or a combination of these mental states, then, can count as blame – provided that they discharge a particular function. Typically, each functional theory of blame identifies a different function of blame. According to some theories, the function of blame is protesting, while others maintain that the function of blame is issuing a response from the agent being blamed. Moreover, it should be noted that most theories of blame allow for cases in which blame discharges more than one function. Functional theories of blame aim at identifying the central function of blame, and not at exhausting every single possible function blame may discharge.

Before presenting the structure of this chapter, let me address a possible concern regarding functional theories of blame. Functional accounts hold that blame discharges an important function in our interpersonal practices. On instrumental interpretations, thus, blame is also *justified* by its function. However, some might argue that the main function of blame is tracking blameworthiness. Thus, blame would be justified by moral reasons –

reasons explaining why agents merit blame – rather than instrumental reasons (Queloz 2020). This is a plausible concern, but I think it conflates two distinct views on the function of blame. The functional accounts of blame I will discuss in this chapter do not attempt to *justify* blame through its function; on the contrary, they *explain* blame by identifying the main function it performs. In other words, not every functional account of blame is an instrumental account of blame. Another way of addressing the same concern is this. The functional accounts of blame that I discuss in this chapter try to find a solution to the problem of the nature of blame; they do not attempt to give an answer to the many questions surrounding the ethics of blame.<sup>41</sup> Now that the issue of instrumentalism has been addressed, let me present the structure of this chapter.

This chapter is divided into three sections. In §1, I present a functional account of blame that puts the notion of protest at its core. In particular, I focus on the theory of blame as protest as developed by Smith. After having presented Smith’s account of blame, I clarify my reasons for rejecting the protesting theory of blame. In §2, I discuss another functional theory of blame. According to such a theory, the main function of blame is to facilitate shared knowledge about the normative changes that have resulted from a wrong (Sliwa 2021). In her 2021, P. Sliwa presents an account of blame according to which the function of blame is eminently epistemic. Even though the epistemic account of blame defended by Sliwa has some appeal, I do not think that it is able to account for the varieties of our blaming practices. Finally, in §3, I examine the functionalist theory of blame recently defended by M. Fricker (Fricker 2016). According to Fricker, blame has the important function of increasing the alignment of the blamer and the wrongdoer’s moral understandings (Fricker 2016, 165). By the end of §3, I clarify why I do not think that the communicative account of blame defended by Fricker is convincing.

---

<sup>41</sup> For reference, see §2, Chapter One.

## §1 Blame as Protest

In this section, I present a functionalist account of blame that puts the idea of protest at the core of our blaming practices. Throughout this section, I refer to such account as the “protesting theory of blame”. According to the protesting theory of blame, the main function of blame is protesting. Before starting my analysis of the protesting theory of blame, let me add an important detail about the methodology that I will follow in this section. If we look at the recent philosophical discussion about blame, we find that there are at least three available versions of the protesting account of blame (Pereboom 2013; Pereboom 2017; Smith 2013; Talbert 2012).<sup>42</sup> In this section, however, I focus on the account of blame defended by Smith. This choice is motivated by the fact that I find her treatment of blame as a form of protest to be the most detailed and convincing. If this version of the protesting account of blame is unable to find a unifying explanation of the varieties of blame, we should be motivated to move to other functional accounts. Now, let us concentrate on Smith’s account to thoroughly understand the matter.

I would like to start by presenting the definition of blame given by Smith:

The Moral Protest Account: To blame another is to judge that she is blameworthy (i.e., to judge that she has attitudes that impair her relations with others) and to modify one own’s attitudes, intentions, and expectations toward that person as a way of protesting (i.e. registering and challenging) the moral claim implicit in her conduct, where such protest seeks some kind of moral acknowledgement on the part of the blameworthy agent and/or on the part of others in the moral community (Smith 2013, 43).

---

<sup>42</sup> P. Hieronymi defends a cognitive account of blame that has tight connection to the protesting account of blame (Hieronymi 2001).



Evidently, Smith elaborates her protesting account of blame from Scanlon's relationship-based account of blame (Scanlon 2008).<sup>43</sup> Smith does not renounce to the cognitive component of Scanlon's account: "To blame another is to judge that she is blameworthy (i.e., to judge that she has attitudes that impair her relations with others)" (Smith 2013, 43). However, Smith modifies the second part of Scanlon's account of blame. Instead of modifying one's intention in a way that reflects the impairment of the relationship, Smith maintains that blame consists in the modification of one's attitude, intentions, and expectations "as a way of protesting (i.e. registering and challenging) the moral claim implicit in her conduct, where such protest seeks some kind of moral acknowledgement on the part of the blameworthy agent and/or on the part of others in the moral community" (Smith 2013, 43). Smith does not hide her debt to Scanlon's account of blame. However, she thinks that the relationship-based account of blame defended by Scanlon does not correctly capture the nuances of our blaming practices.<sup>44</sup> Before presenting the details of Smith's account of blame, let me address a possible concern about my treatment of this particular account of blame. In the previous chapter, I classified Scanlon's account of blame as a conative theory, which is a kind of *content-based theory* of blame. However, in this section I present Smith's theory as a functional account of blame. If Smith builds her account of blame on Scanlon's, should not we classify both their theories as pertaining to the same family of theories? In other words, if we considered Scanlon's account as a content-based account of blame, should we not do the same with Smith's? Additionally, note that the opposite could also be true: if we want to classify Smith's account as functional, should we not also classify Scanlon's account as a functional theory of blame? These are legitimate worries, but I have two ways of addressing these concerns. First, I am not alone in classifying the protesting account of blame as a functional account of blame, while at the same time treating Scanlon's account as

---

<sup>43</sup> For reference, see §4, Chapter Two of this dissertation.

<sup>44</sup> Smith also discusses Sher's account of blame (Smith 2013). However, she seems to draw much of her ideas about blame from the relationship-based account defended by Scanlon.

being content-based. Coates and Tognazzini offer the same taxonomy of these theories (Coates and Tognazzini 2021). Second, even if Smith presents her theory as an improvement of Scanlon's account of blame, the protesting account of blame can plausibly be read as a functional account. According to such an account, the function of blame is protesting the moral claim implicit in the conduct of the person being blamed. After all, Smith explicitly claims that blame does not simply consist in the modification of one's attitude, intentions, and expectations; instead, she adds that blame consists in such modifications *as a way of protesting* (Smith 2013, 43). Thus, it is not implausible to classify her account as a functional account of blame.

Putting these worries aside, it is now time to critically evaluate Smith's proposal. Is her account capable of addressing the most problematic aspects of other theories of blame? Let me start by mentioning the merits of the protesting account of blame. First, as Chislenko also notes, this account seems to capture two important shared intuitions about blame (Chislenko 2019, 166). Many authors who work on blame identify blame as an inherently (or, at least, in an important way) communicative phenomenon (Fricker 2016; McKenna 2012; Shoemaker 2015). Surely, blame is not always outwardly expressed.<sup>45</sup> Even so, many take blame as connected to the idea of addressing the person being blamed. The second feature of blame that the protesting account seems able to capture is the connection between blame and an element of resistance. According to this line of thought, to blame someone is to resist some element of her behavior. We might understand the protesting account of blame as combining these two intuitions about blame. To put it in a different manner, *blame communicates resistance*.

---

<sup>45</sup> We can blame each other even if we do not *express* our blame, e.g., in cases of what I have called "private blame". See §5, Chapter One.

Smith identifies at least three other important advantages of her theory over the relationship-based account of blame developed by Scanlon and other content-based theories of blame. First, she claims that “this account gives a clear explanation of how blaming someone goes beyond simply judging her to be blameworthy” (Smith 2013, 43). According to the definition of blame given by Smith, blame requires two distinct components. The first element is a judgment of blameworthiness, while the second component is a modification of one own’s attitudes, intentions, and expectations towards an agent as a way of protesting the moral claim implicit in that agent’s conduct. Being able to explain in a straightforward way how blaming someone goes beyond the mere judgment of blameworthiness is surely an advantage over cognitive accounts of blame, which struggle to identify a criterion to distinguish judgments of blameworthiness and instances of active blame.<sup>46</sup> However, most of the other content-based accounts of blame have the same advantage over cognitive accounts. So, Smith has to provide additional reasons to make us embrace her protesting account of blame.

One of the most important features of Smith’s protesting theory of blame is that it “makes clear why (blame) is a response that it is appropriate to direct only at individuals we regard as morally responsible for their conduct” (Smith 2013, 46). When we blame someone, according to the definition of blame proposed by Smith, we are protesting the moral claim implicit in the conduct of the blamed agent. If this is so, then we treat that agent as someone who is able to make moral claims, as someone whose conduct can be rightly challenged. In other words, blaming someone is a way of treating the blamed agent as a normatively competent agent. Surely, this is an important component of our intuitive concept of blame and, more importantly, of our blaming practices. However, it seems to me that many content-based theories of blame can capture the very same intuition. Consider Strawson’s account of

---

<sup>46</sup> For a discussion, see §1, Chapter Two.

reactive attitudes (Strawson 2008). According to this account, an agent is blameworthy when she is an apt target of reactive attitudes, while blame consists in actively targeting an agent with said reactive attitudes. In turn, blame is appropriate when it targets blameworthy agents. Remember that most emotional theories of blame take their cue from Strawson's influential account. Then, it seems that emotional theories of blame are able to show why blame is a response that is appropriate to direct at individuals we regard as (morally) responsible. If this is so, then, the protesting account of blame defended by Smith is not the only available theory of blame which explains in a straightforward way why blame is a response that is appropriate to direct only at individuals we regard as responsible for their conduct.

The protesting account of blame has at least two more positive aspects. Smith argues that the protesting account of blame is able to find a unifying explanation of the diversities of blame. According to Smith, the function of blame is protesting the moral claim implicit in someone's conduct. Blamers may protest the moral claim implicit in the conduct of the person being blamed even without feeling any blaming emotion. Even if emotions may play an important role in our blaming interactions, they are not a necessary component of the protesting account of blame. Thus, Smith's account of blame has a straightforward explanation of cases of dispassionate blame. Moreover, the protesting account of blame is also able to account for cases of blaming the absent and cases of blaming the dead. We can protest the moral claim implicit in Nero's or Caligula's violent conduct even if Roman emperors passed away long ago. Clearly, our protests directed towards the dead do not aim at issuing any apology from the person being blamed. Still, in such cases protest "can have as a secondary aim moral recognition on the part of the wider moral community" (Smith 2013, 44). To protest the moral claim implicit in the conduct of an historical figure means *to communicate to the living* (the "wider moral community") that certain behaviors are still not to be tolerated. Before moving to the discussion of the protesting account of blame, let me

quote Smith one more time. According to Smith, the protesting account of blame “does not seem to allow ‘false positives’, [...] in which the basic conditions of blame are met but we are disinclined to classify the resulting reactions as instances of blame” (Smith 2013, 46). As I will show in the following paragraphs of this section, however, I do not think that Smith is right. Not only does her account allow for false positives, but it is also unable to give a convincing explanation of a significative portion of our blaming interactions. Moreover, the protesting account of blame runs the risk of being circular.

The first issue of the protesting theory of blame is the following. How should we interpret the meaning of the core notion of the protesting account of blame? In other words, what use of the term “protest” is made by the protesting account of blame defended by Smith (Chislenko 2019)? The protesting account of blame can use the term “protest” in two ways, either the ordinary meaning of the term or the technical significance of it. In the first case, the protesting account of blame seems unable to account for cases of private blame, and thus runs into counterexamples. In the second case, at least according to Smith’s version of the protesting account of blame, the concept of protest seems dependent on the concept of blame – thus, the protesting account of blame would run the risk of being circular.

Let us start by considering the first case. Hypothesize that the protesting account of blame uses the ordinary notion of protest. According to our ordinary language, a protest is an outwardly expressed communicative act. Workers can go on a strike and protest outside of their offices, citizens can protest the behavior of the politicians ruling their nation, and the members of the crew of a ship can protest the conduct of their captain. These all are clear examples of protest which involve an outwardly expressed communicative act. Is it plausible to think that the protesting account of blame uses this ordinary notion of protest? I think that the answer to this question must be negative. To see why, consider cases of private blame, e.g., blame that is not outwardly expressed. How could the protesting theory

of blame account for cases of private blame by using the ordinary notion of blame? In other words, can there be something like a “private protest”? In presenting her theory of blame, Smith acknowledges that private blame could be a problem for her protesting account of blame:

“But we can also protest ill treatment privately through the modification of other attitudes, intentions, and expectations. Even if we are not in a position (for whatever reason) to make these attitudinal modifications known, I believe these reactions embody, at a deep level, both moral protest and a desire that the wrongdoer morally acknowledge his wrongdoing. Blame is incipiently communicative both in the sense that it registers (i.e., communicates, even if only to the victim herself) the existence of unjustified wrongdoing and in the sense that it seeks some sort of moral recognition of wrongdoing on the part of the blameworthy agent” (Smith 2013, 44).

In my opinion, this passage is not very helpful for clarifying the notion of private protest. Admittedly, blamers can register a wrong by communicating only with themselves. However, this does not seem to be close to our ordinary use of the term “protest”. Moreover, how could blame that is not even expressed seek “some sort of moral recognition of wrongdoing on the part of the blameworthy agent” (Smith 2013, 44)? It seems to me that the treatment Smith reserves for private blame is not very convincing. Proponents of the protesting account of blame, however, could refer to the notion of “communicative entity” in order to defend their theory (Macnamara 2015a; Macnamara 2015b). Macnamara enumerates emails and letters as communicative entities. Letters and emails are communicative entities even when they do not actually perform their communicative activity. After all, a letter can be kept in the drawer of your desk, without being sent to its addressee. Similarly, an email can be kept as a draft and not being sent. In other words, communicative entities need an additional

human intervention in order to perform their communicative function. Nonetheless, they do not stop being communicative entities even when their function is not carried out with the help of any human interactions (Macnamara 2015b, 217). A proponent of the protesting account of blame could think that the notion of communicative entity can help account for cases of private blame. Blamers can modify their attitudes, intentions, and expectations even without letting others know about these modifications.<sup>47</sup> So, it can be argued that these modifications of attitudes, intentions, and expectations act just like the communicative entities described by Macnamara. After all, the notion of “private protest” seems intelligible (Chislenko 2019, 169). But does this resonate well with how Smith characterizes her account of blame as protest? I think it does not.

In the first part of this section, I have quoted the definition Smith gives of her protesting account of blame. For my present purpose, we can put aside the first part of this definition, according to which blame requires a cognitive element describable as a judgment of blameworthiness. Let us focus on the second part of Smith’s treatment of blame. According to this second part of the definition of the protesting theory, blame consists in the modification of one own’s attitudes, intentions, and expectations toward the person being blamed *as a way of protesting* the moral claim implicit in her conduct (Smith 2013, 43). If this is so, however, it is hard to think that Smith can allow the blamer’s modification of attitudes, intentions, and expectations to be considered as communicative *entities*. In treating these modifications as ways of protesting, Smith seems to refer to the ordinary notion of “protest”. According to the ordinary meaning of “protest”, however, protest is not considered to be a communicative *entity*, but a communicative *activity*. Moreover, as I have noted above, in addressing the concerns about private blame Smith also adds that “blame is incipiently communicative [...] in the sense that it registers (i.e., communicates, *even only to*

---

<sup>47</sup> Smith herself could be interpreted as suggesting something along these lines in the quoted passage (Smith 2013, 44).

*the victim herself*)” (Smith 2013, 44, emphasis added). As Chislenko also notes, if Smith were not considering protest as a communicative *activity* rather than a communicative *entity*, the caveat “even only to the victim herself” would be unnecessary (Chislenko 2019, 169). By adding this caveat, Smith strongly suggests that the protesting account of blame employs the term “protest” in its ordinary use. If this is true, however, the solution offered by Macnamara is no longer helpful. While communicative entities can be kept private and still be communicative, the concept of a private communicative activity seems unintelligible. Thus, the protesting account of blame seems unable to account for cases of private blame. Is this sufficient to dismiss the protesting account of blame? Maybe it is not. What if the protesting account of blame employs “protest” as a technical term, instead of referring to the ordinary use of protest?

So, let us consider whether another interpretation of the term “protest” could help the protesting account of blame face the challenge posed by private blame. Obviously, an alternative interpretation of protest cannot rely too heavily on our ordinary use of the term “protest” – otherwise, it would run into the same problem that I have highlighted above. At the same time, the alternative interpretation of protest that we are looking for cannot depend on the definition of blame, since it is supposed to *define* blame, and not the other way round. Otherwise, the protesting account of blame runs the risk of being circular. Smith herself suggests a seemingly promising way to avoid relying too heavily on the ordinary use of the term “protest”. She does so by suggesting that blamers do not protest *persons*, but *the moral claim* implicit in the behavior of the person being blamed. Instead of addressing a person, blamers address a moral claim implicit in the conduct of that person. This seems to avoid the problems faced by the ordinary use of the term “protest”. If this is correct, it seems that Smith employs protest in a broader sense than the ordinary one. And, as Chislenko aptly notes, a broader notion of protest might include emotional reactions, unexpressed reactions,



and even reactions that do not seek acknowledgment by the person being blamed (Chislenko 2019, 173-176). After all, the function of blame is to challenge the moral claim implicit in the behavior of an agent: blame protests a claim, not a person. Thus, even the notion of inner protest becomes intelligible. The function of protest is to challenge a moral claim; in turn, a protest is still a protest even if it is not outwardly expressed. Thus, it seems that Smith's broader notion of protest can avoid the problems faced by a narrow interpretation of the term "protest". However, I think that this broader notion of protest as a means of challenging a moral claim faces another serious problem. According to Smith, blame consists in a judgment of blameworthiness together with a modification of one's attitudes, intentions, and expectations *as a way of protesting a moral claim*. As I have just argued, the protesting account of blame needs to employ a broader notion of protest in order to avoid counterexamples like cases of private blame. According to this broader notion, protest *consists in challenging a moral claim*. If this is correct, the protesting account of blame defended by Smith runs the risk of being circular, since it seems that the notion of protest is dependent from the notion of blame. The notion of protest as a way of challenging a moral claim, however, should *define* blame and its function. In other words, it seems to me that in the protesting account defended by Smith blame acts both as the *definiens* and the *definiendum*. Finally, even if there is a way out of circularity, the protesting account seems at best uninformative. As Coates and Tognazzini also notes, "it's not clear that protest is independent of blame, such that one could specify what it is to protest without appealing to blaming attitudes" (Coates and Tognazzini 2021, §1.4). After all, it may be true that blame is a form of protest, provided that we define protest in a new, broader sense. However, without an independent definition of protest, this proposition remains uninformative. Thus, it can be concluded that Smith's protesting account of blame faces a serious challenge even if it employs a technical term of protest.

In my view, the protesting theory of blame can either use an ordinary notion of “protest”, or it can employ the term “protest” as a technical term. However, I have shown that in both cases the protesting account of blame faces important challenges. On a narrow, ordinary interpretation of protest, the protesting theory of blame is unable to account for cases of private blame. On a broader, technical interpretation of protest, on the other hand, the protesting account of blame runs the risk of being circular – and even if it avoids circularity, it loses almost all of its explanatory power. This is not the only problematic aspect of the protesting account of blame. As I have mentioned above, Smith thinks that one of the advantages of her theories over other accounts of blame is that of not allowing cases of false positives, e.g., cases in which “the basic conditions of blame are met but we are disinclined to classify the resulting reactions as instances of blame” (Smith 2013, 46). However, I think that Smith’s account actually allows for false positives. It seems to me that most of us can protest the moral claim implicit in someone’s conduct even without blaming that person, and that this can be done in conjunction with a judgment of blameworthiness. Parents may judge their children to be blameworthy, and they can modify their attitudes as a way of protesting the moral claim implicit in the behavior of their children.<sup>48</sup> However, this does not necessarily amount to blame. More generally, it is plausible to think that claims – be they implicit or explicit – can be challenged in many ways other than blame. We can challenge the claim implicit in the conduct of other agents with sorrow, contempt, sadness, horror, and even by ignoring the agent whose moral claim we are protesting.<sup>49</sup> In other words, modifying one’s attitudes, intentions, and expectations as a way of challenging a moral claim may not be enough to blame. Thus, it seems that the protesting account of blame does allow for false positives.

---

<sup>48</sup> That is, provided the children are not too young to be the kind of creatures who are able to make moral claims – even if just implicitly.

<sup>49</sup> In cases of self-blame, regret may be a way of challenging the claim implicit in our own conduct. However, it is not clear that when we feel regret, we are blaming ourselves. For a discussion, see Portmore 2022.

Before moving to the next section, I want to note one additional problematic aspect of the protesting account of blame. In Chapter One of this dissertation, I have argued that a satisfying account of blame should help us understand cases of non-moral blame.<sup>50</sup> The protesting account of blame that I have presented in this section, however, does not seem well-equipped to explain cases of non-moral blame. According to Smith, the function of blame is protesting the moral claim implicit in the behavior of the person being blamed. Smith does not seem interested in expanding her account of blame to cases of non-moral blame.<sup>51</sup> However, for the sake of the argument, we could try to apply Smith's account of blame to cases of non-moral blame. One possible way in which Smith's account can be adapted to explain cases of non-moral blame is the following. Instead of claiming that blame protest the *moral* claim implicit in the behavior of the person being blamed, an alternative version of the protesting account of blame could argue that blame protests the *normative* claim implicit in the behavior of the person being blamed. This way, the protesting account of blame could attempt to explain cases of non-moral blame, e.g., cases in which a person is being blamed for the normatively inadequate (but not necessarily morally problematic) claim implicit in her conduct. This seems the most natural way to modify the protesting account of blame in order to explain cases of non-moral blame. However, while the idea of "moral protest" seems plausible, some would find the idea of a "non-moral protest" perplexing. In cases of non-morally faulty actions, protest may not be a fitting reaction. Moreover, it could be objected that the very idea of a "normative claim" implicit in one's behavior is not entirely intelligible. Obviously, Smith could address these concerns by claiming that blame is an inappropriate reaction in cases of non-morally faulty behavior. However, at least at an intuitive level, our interpersonal practices do include cases of non-moral blame. Thus, Smith should at least provide a reason to exclude cases of non-moral blame from her protesting

---

<sup>50</sup> See §4, Chapter One.

<sup>51</sup> In her treatment of blame as a form of protest, Smith never mentions non-moral blame.

account of blame. Absent these reasons, it seems plausible to conclude that the protesting account of blame may not be our best option in order to explain cases of non-moral blame. This is not a knock-down argument against the protesting account of blame, but it can motivate us to look for other functional accounts of blame.

In the following sections, I turn to another recent functional account of blame defended by P. Sliwa.

## §2 The Epistemic Function of Blame

In the previous section, I have argued that the protesting theory of blame presents some relevant problems. However, I think that the theory of blame defended by Smith highlights significant details about our blaming practices. The role of blame may not be that of protesting, but it nonetheless may be that of facilitating shared knowledge among members of a community. In this section, I consider whether this can be the main function of blame. In order to investigate this possible solution to the problem of the nature of blame, I discuss the account of blame recently defended by P. Sliwa. According to Sliwa, the function of blame is precisely that of facilitating shared knowledge among members of the same community (Sliwa 2021). Before proceeding, I want to highlight an important feature of the account of blame presented by Sliwa. Unlike Smith, Sliwa does not present her theory of blame as an improved version of another content-based account of blame.<sup>52</sup> Before presenting Sliwa's epistemic theory of blame, let me highlight this passage: "To investigate its [of blame] nature, I draw on a methodology that has been popular in epistemology and attend to the function of blame" (Sliwa 2021, 200).

---

<sup>52</sup> As I have showed in the previous section, Smith presents her account of blame as an improved version of Scanlon's theory of blame.

According to Sliwa, “the function of blame is to facilitate shared knowledge about the normative footprint – that is, the normative changes that have resulted from a wrong” (Sliwa 2021, 203). If this is true, then the function served by blame is eminently epistemic: wrongdoings leave a normative footprint, and blame lets us share knowledge over this fact. I find this view interesting and worth discussing. It may well turn out that the function of blame is indeed epistemic. However, I find that the characterization of blame as an epistemic practice offered by Sliwa is unconvincing. By the end of this section, I will give some reasons why I think the epistemic account of blame defended by Sliwa should be rejected. For now, let me focus on an analysis of Sliwa’s account of blame and her justification for this theory.

First, Sliwa denies that content-based theories of blame can give a unifying explanation of our diverse blaming practices. Sliwa examines our blaming interactions, explicitly claiming that our blaming practices are incipiently communicative practices (Sliwa 2021, 202). Sliwa follows McKenna in claiming that when we blame someone, we *start a conversation* with the person being blamed (McKenna 2012). As a form of linguistic act, blame requires an uptake from the person being blamed, e.g., asking for forgiveness or accepting the blame and taking responsibility. Moreover, Sliwa acknowledges the existence of dispassionate blame: not all blamers need to feel what some authors call “the blaming emotion”.<sup>53</sup> Blaming practices are so diverse that content-based theories of blame cannot account for all the different ways in which we blame each other. I take this to be an advantage of the epistemic account of blame over content-based theories of blame.

Then, Sliwa moves on to characterize the notions of wrongdoing and of normative footprint, which play an important role in her account of blame (Sliwa 2021, 203-205). A

---

<sup>53</sup> Sliwa explicitly takes into account cases of what I have called “dispassionate blame”: “We do often blame angrily and resentfully. But we also blame without feeling or expressing any affect at all: simply by noting that a misstep was done, and an apology is in order. Equally, we blame with bemusement or even Schadenfreude – particularly, in gossip blame” (Sliwa 2021, 202).

wrongdoing occurs when an agent performs an action that violates some moral obligations she is under, or when she violates someone else's moral right. Notice that Sliwa focuses on *moral* wrongdoing, taking into account violations of moral obligations and violations of moral rights. I think that restricting the analysis to moral blame, at least partially, clashes with the premise of focusing on our actual blaming practices. After all, it seems to me that our blaming practices include instances of non-moral blame, too. I will come back to this point later. For now, let me continue to describe the account of blame Sliwa puts forward. Plausibly, moral wrongdoing can cause harm and even pain. However, Sliwa focuses on the normative consequences of wrongdoing. Wrongdoing alters the normative landscape constituted by the web of rights, duties, and permissions in which we are all embedded (Sliwa 2019, 201). Sliwa calls this alteration of the normative landscape brought up by wrongdoing the “normative footprint” of wrongdoing. In turn, this normative footprint can be constituted by three distinct elements: the creation of reparative rights and duties; the alteration of feeling norms; the alteration of relationships norms.<sup>54</sup> Wrongdoing alters the normative landscape by giving rise to duties, such as the duty to recognize one own’s wrong actions. Moreover, the wronged person often acquires the right of receiving an apology from the wrongdoer. The normative footprint left by wrongdoing may also alter norms about how we should feel towards each other. For example, a person may be entitled to feel angry or resentful after having been wronged. Finally, wrongdoing may alter the normative landscape by modifying relationship norms. Friends accept that their relationship involves certain norms and special duties. Wronging a friend can alter the normative landscape by altering these norms and these special duties. Friends usually remember each other’s birthdays, and they sometimes send cards to celebrate. But if I repeatedly wrong a friend of mine, it could be permissible for my friend to stop sending cards for my birthday. To sum up, according to Sliwa wrongdoing changes the

---

<sup>54</sup> Differently from the exercise of normative powers, which alter the normative landscape by voluntary actions performed by agents, wrongdoing alters the normative landscape in ways that are usually unwelcomed by the wrongdoer (Owen 2012).

normative landscape – defined as the web of duties, rights, and permissions we are all embedded in – by leaving a normative footprint. In order to critically evaluate the functional account of blame defended by Sliwa, I should add one more element to the characterization of the concepts of normative landscape and of normative footprint.

Agents who are embedded in the web of duties, rights, and permissions which constitute the normative landscape typically care about knowing “the normative lay of the land” (Sliwa 2021, 205). Agents are interested in knowing how to discharge their obligations towards each other, and they care about avoiding wrongdoing. Agents who are interested in knowing the “normative lay of the land” also care about the alterations of the normative landscape. And, as I have shown earlier, wrongdoing alters the normative landscape by leaving a normative footprint. So, it seems plausible that agents embedded in the normative landscape are also interested in engaging in practices that make normative footprints public. Among these practices, blame has a prominent position. According to Sliwa, blame has the important function of informing others about the changes in the normative landscape, alerting both the wrongdoer and other members of the community about the normative footprint left by the wrongdoer’s action. This communicative role of blame is taken to be of great relevance by Sliwa: blame is not only important because the normative footprint left by wrongdoing can be serious, but also because sometimes wrongdoers are not even aware of the normative consequences of their actions – as when agents wrong others inadvertently or involuntarily (Sliwa 2021, 206).

Before moving to the problematic aspects of this epistemic account of blame, let me pause for a moment in order to mention two clarificatory points made by Sliwa (Sliwa 2019, 207). First, Sliwa states that her account does not aim at explaining every function that blame may serve. Blame may have many functions. However, her account aims at identifying the *main function* of blame. And that function, as I have shown, is to facilitate the shared

knowledge about the normative footprint left by wrongdoing. I think that Sliwa is correct in claiming that her account of blame aims at identifying the core function of blame and of our blaming practices. However, I consider this as a distinctive feature of every functional account of blame. Functional accounts of blame do not attempt to explain every possible function that blame may discharge. Functional accounts aim at identifying the core function of blame. So, even if Sliwa is correct, I do not think that adding this clarificatory point adds much to her proposal – after all, it is a feature that her account shares with every other functional theory of blame. Finally, Sliwa adds that “to say that the function or point of blame is to facilitate shared knowledge is to make a claim about the constitutive aim of the practice as a whole. It is to say that this aim is (partly) what makes our blaming practice a blaming practice. It is not a claim about each individual blaming interaction” (Sliwa 2019, 207). This means that some blaming interactions may not discharge the epistemic function described by Sliwa. In turn, it is also implied that discharging the epistemic function is not a necessary condition for a given interaction to count as a blaming interaction. Thus, referring to counterexamples in which the epistemic function is not discharged by the blaming practice will not constitute a threat against Sliwa’s account of blame. This is an important feature of the epistemic account of blame defended by Sliwa, since our actual blaming practices play a prominent role in it.

I now turn to the discussion of the epistemic account of blame that I have presented in this section. My criticism against the theory of blame defended by Sliwa is articulated into two main points. First, I think that the epistemic account of blame cannot account for some data points that I have identified in the first chapter. Second, the epistemic account of blame seems inadequate to account for cases of non-moral blame. However, our actual interpersonal practices include cases of non-moral blame. Given the attention Sliwa reserves



for our interpersonal practices, one would expect her account to be able to explain cases of non-moral blame.

Let me start by putting into test the epistemic account of blame against some of the data points I have listed in §5, Chapter One. The epistemic account of blame has a straightforward explanation of cases of third-party blame. According to Sliwa, “we would expect a blaming practice to encompass second-personal as well as third-personal varieties: blaming wrongdoers “to their face”, blaming them in their absence to third parties, and blaming others even when their wrongdoing has not affected us personally at all. On the epistemic function account, we would thus expect blame to come in many second- and third-personal varieties” (Sliwa 2021, 208). Blame serves its epistemic function not only because it facilitates shared knowledge between the blamer and the person being blamed, but also because it can reach a wider audience. I take this to be an advantage of the epistemic account of blame over other theories that focus primarily on directed dyadic blame. The epistemic account of blame, however, faces an important challenge when it has to explain cases of private blame. Even at an intuitive level, it can be objected that in order to facilitate shared knowledge about something blamers must express their blame. If this is true, then unexpressed blame cannot discharge its main function, e.g., facilitate shared knowledge about the normative footprint left by wrongdoing. Sliwa directly addresses this potential issue of her account of blame. She argues that “for many interpersonal practices, there is an intra-personal correlate. Chess is an interpersonal practice but I can play chess by (“against”) myself [...] Thus, we should not be surprised to find a private correlate to public blame. We blame others privately by thinking the things we would say “out loud” were we to blame them publicly” (Sliwa 2021, 209). Is this sufficient to account for cases of private blame? I do not think it is. First, I do not think that we blame others privately by thinking the things we would say “out loud”, as Sliwa argues. On the contrary, there are many possible scenarios

in which we would *not* say “out loud” what we think, even if the person being blamed were present. Most of the time, we have prudential reasons to avoid publicly expressing our blame. Thus, I do not think that private blame can be considered as the exact correlate to public blame. Moreover, how could unexpressed blame increase our knowledge about the normative footprint left by wrongdoing? Consider this passage: “By thinking through the accusations and demands – articulating them to myself – I can come to a better understanding of what it is that the wrongdoer has done and owes as a result. And in this way, private blame can facilitate shared knowledge indirectly” (Sliwa 2021, 209). Sliwa argues that private blame facilitates shared knowledge indirectly because the blamer “thinks through the accusations and demands” by articulating them to herself. However, this does not seem convincing. Consider an example in which an agent *cannot* outwardly express her blame, whether because she is physically unable to do so, or because she has overriding prudential reasons not to express her blame. In such circumstances, I strongly doubt that the blamer would “think thorough” the accusations and demands, thus indirectly facilitating shared knowledge about the normative footprint left by wrongdoing. I find it much more plausible that the blamer would react with anger, resentment, indignation, or in an otherwise passionate way.<sup>55</sup> And this seems far from being an occasion for the blamer to articulate the accusations and demands to herself. Thus, I think that Sliwa’s treatment of cases of private blame is unconvincing. As a result, we have reasons to think that the epistemic account of blame faces a challenge when it comes to addressing the data points of blame that I have identified in §5, Chapter One.

The epistemic account of blame presented by Sliwa has another important problem. Sliwa focuses on our actual practices in order to investigate blame. Instead of focusing on the conceptual analysis of blame, she considers blame as a “social phenomenon”. Thus, one

---

<sup>55</sup> It could also be argued that an agent would react in such ways *precisely because* she is unable to outwardly express her blame.

would expect that Sliwa's theory of blame should be able to explain the majority of our actual practices. But as I have noted in §4, Chapter One, our actual practices seem to include cases of non-moral blame along with cases of moral blame. However, Sliwa does not even mention non-moral blame, despite her interest in our actual practices. My contention is that the epistemic account of blame is unable to account for cases of non-moral blame. To understand this point, consider first Sliwa's treatment of the distinction between judgments of blameworthiness and instances of active blame. According to Sliwa, our actual interpersonal practices *do not lend support* to the distinction between judgments of blameworthiness and instances of active blame: "the epistemic function account is right not to distinguish between a blaming practice and a practice of making and communicating judgments of blameworthiness. This is because our moral practice lends no support to such a distinction" (Sliwa 2021, 210). Consider this exchange between two friends:

A: You did not show up on time at our last *rendezvous*. I was waiting for you!

B: I am not sure whether you are blaming me, or merely judging me blameworthy. Could you be more precise?

A: What are you saying?!

It may well be that the distinction between judgments of blameworthiness and blame plays an important role at the conceptual level. However, our actual practices do not lend support to this distinction; thus, the epistemic account of blame is right in not distinguishing between judgment of blameworthiness and blame (Sliwa 2021, 210-211). Given the attention Sliwa gives to our actual interpersonal practices and given that our interpersonal practices include cases of non-moral blame, one would expect the epistemic account of blame to be able to account for cases of non-moral blame. As I have anticipated, however, I do not think that the epistemic account of blame is well-equipped to account for non-moral blame. According to Sliwa, the main function of blame is to facilitate shared knowledge among members of the same community about

the normative footprint left by wrongdoing. This is explained by our interest – as human beings – in knowing the “lay of the [normative] land” (Sliwa 2021, 205). However, can we plausibly say that the chess player who realizes his blunder blames himself in order to promote an interest in knowing the changes in the normative landscape? Or consider a case of epistemic blame. When we say that someone who believes to possess magical powers makes poor use of his intellectual faculty, are we facilitating shared knowledge about the normative footprint left by his wrongdoing?<sup>56</sup> The epistemic account of blame does not seem to be our best resource to account for cases of non-moral blame. Maybe Sliwa has an explanation for such cases, but I suspect that the epistemic account of blame would need to be modified. As it is defended by Sliwa in her 2021, I do not think that the epistemic account of blame is able to account for cases of non-moral blame. Thus, we are motivated to look for another theory of the nature of blame.

In the next section, I will present and discuss an alternative functional theory of blame.

### **§3 Communicative Blame: A Paradigm for our Blaming Practices?**

The functional accounts of blame that I have discussed in the previous two chapters present some similarities. It can plausibly be argued that both the protesting theory of blame and the epistemic account of blame highlight the communicative dimension of blame. What if we focus on this communicative dimension, and we make *it* the core function of blame?

In her 2016, M. Fricker proposes a functional theory of blame that focuses on the communicative dimension of blame (Fricker 2016). Before presenting the details of her

---

<sup>56</sup> Moreover, one could also ask *if* this is a case of *wrongdoing*. We should consider whether performing a *wrong* action is a necessary condition for blame.

theory, let me highlight the methodology employed by Fricker. In order to understand blame and our blaming practices, Fricker maintains that we should adopt a *paradigm-based* account of blame: “by constructing a candidate paradigm—a form of blame I shall label ‘Communicative Blame’— I hope to build a platform from which I can account for non-paradigmatic cases as derivative” (Fricker 2016, 166). Fricker adopts this paradigm-based approach as an answer to the manifest diversity of blame. The traditional philosophical analysis – intended as the attempt to describe necessary and sufficient conditions – 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” (Fricker 2021, 166). The failure of content-based theories of blame to offer a satisfying explanation of blame reflects this difficulty. In order to overcome such difficulty, Fricker opts for a different kind of analysis. Instead of relying on the instruments of conceptual analysis, she moves to the study of our blaming practice, trying to identify the most simple and straightforward form of such practice – that is, the paradigm of blame.

Fricker identifies the paradigmatic form of blame in what she labels as “Communicative Blame”. Communicative Blame is defined as “blame that is performed in the most simple and socially immediate sort of interpersonal exchange: I wrong you, and in response you let me know with feeling that I am at fault for it. It is an essentially second-personal, I-Thou interaction” (Fricker 2021, 171). Communicative Blame is not necessarily verbal – after all, human beings frequently communicate while being silent. A dismissing gaze or a cold handshake, for example, may communicate contempt or resentment to the recipient of such non-verbal activity. According to Fricker, what distinguishes blame from other reactions to wrongdoing such as shock, disgust, or sorrow is the fact that “in Communicative Blame the blamer *accuses the wrongdoer of fault?*” (Fricker 2021, 171). In turn, the expression of

this judgment of fault is usually accompanied by the feeling of an emotion such as rage or resentment on the part of the blamer. According to Fricker, once we have sufficiently characterized this paradigmatic form of blame, we will be able to explain all the other instances of blame by considering them as derivative.

There is an important element that is part of the characterization of the paradigmatic form of blame that is still missing. So far, I have not mentioned the *point* or the *function* of Communicative Blame. In order to identify its function, Fricker enquires from what sort of speech act Communicative Blame gains verbal expression (Fricker 2016, 172). It is plausible to categorize Communicative Blame as an *illocutionary* speech act, which cannot completely discharge its function without an uptake from the hearer.<sup>57</sup> Being an illocutionary speech act, Communicative Blame can be fully understood only by enquiring which is its illocutionary point. What is the illocutionary point of accusing someone of fault?<sup>58</sup> According to Fricker, Communicative Blame does not merely serve an epistemic function, e.g., let the wrongdoer know that she has violated a norm. Moreover, accusing someone of fault does not only aim to make the recipient of the accusation feel bad for what she has done. Fricker argues that the illocutionary point of Communicative Blame – its function – is “to make the wrongdoer feel *sorry for what they have done*” (Fricker 2016, 172). More precisely, Communicative Blame aims at instilling *remorse* in its recipient, where remorse is understood as a moral emotion. Remorse, moreover, is also constituted by a cognitive component: in feeling remorse, the wrongdoer painfully understands the wrong she has committed. Communicative Blame presents an analogous structure: “[it is] a cognitively loaded moral emotion, this time a perception of a wrong one suffers at the hands of another” (Fricker 2016, 173). Remorse and blame work together in order to let Communicative Blame discharge its primary

---

<sup>57</sup> The traditional analysis of speech acts was provided by J.L. Austin in his *How to Do Things with Words* (Austin 1962).

<sup>58</sup> Recall that, according to Fricker, blame is distinct from other reactions to wrongdoing because blamers accuse the wrongdoer of fault (Fricker 2021, 171).

function: increasing the alignment of moral understanding between the blamer and the person being blamed. Ideally, the function of blame is successfully discharged when the moral understanding of the person being blamed and the moral understanding of the blamer converge. Admittedly, in real-life scenarios the moral understanding of the blamer and the moral understanding of the person being blamed rarely converge; however, blame in part discharges its function by triggering a “moral conversation” between the blamer and the person being blamed.<sup>59</sup>

Finally, Fricker also enquires which perlocutionary function is served by Communicative Blame. Given that no one likes being blamed and no one enjoys being found fault with, the perlocutionary point of blame is “manifest [...], (it is) to prompt a change for the better in the behaviour (inner and outer) of the wrongdoer” (Fricker 2016, 173). Wrongdoers who get blamed have a chance at increasing their own moral awareness; in turn, this increased moral awareness may prompt a change for the better in the conduct of wrongdoers. Thus, Fricker has a straightforward explanation of the “proleptic mechanism” of blame.<sup>60</sup> By increasing the moral awareness of the person being blamed, blame prompts a change in the conduct of the person being blamed. Blame does not only accuse the wrongdoer of having performed a faulty action in the past, but it also influences the wrongdoer’s future conduct.

Admittedly, the treatment of blame offered by Fricker is fascinating and worth discussing. One of its merits, as I have just argued, is identifying the proleptic function of blame – a feature of blame that many other theories fail to consider. However, I do not think that the paradigm-based account of blame defended by Fricker offers a solution to the problem of the nature of blame. First, we should note that Fricker considers what she calls

---

<sup>59</sup> In this respect, the paradigm-based account defended by Fricker draws from the conversational model of responsibility developed by Mckenna. For reference, see McKenna 2012.

<sup>60</sup> For the proleptic mechanism of blame, see Williams 1995.

“Communicative Blame” the most basic – paradigmatic, indeed – form of blaming interaction. However, as I have noted multiple times in this dissertation, our blaming activities include different ways in which we blame each other. Why not consider *private* blame the paradigmatic form of blame that explains the other instances of blame as derivative?<sup>61</sup> Alternatively, it could be argued that *self-blame* is the paradigmatic form of blame. Admittedly, it would be difficult to maintain that the function of, say, self-blame is increasing the alignment of moral understanding between the blamer and the person being blamed. After all, the blamer and the person being blamed are the same agent. Clearly, this is not a knockdown argument against Communicative Blame as the paradigmatic form of blame. However, I do not think that Fricker offers sufficient reasons to maintain that what I have called directed dyadic blame<sup>62</sup> is the most basic kind of blame.

Second, the paradigm-based account of blame defended by Fricker intuitively inherits the most problematic aspect of communicative theories of blame. How should the paradigm-based theory deal with cases of private blame, e.g., blame that is not outwardly expressed? Let us recall that the main function of blame is increasing the alignment of moral understanding between the blamer and the person being blamed. In turn, the perlocutionary point of blame is prompt a change for the better in the conduct of the person being blamed. However, it could be argued that unexpressed blame cannot discharge neither the main function of blame, nor its illocutionary force. How can private blame influence the wrongdoer’s conduct? Note that the same point can be made about blaming the distant or blaming the dead. It is hard to understand how blame that never reaches its recipient could influence the wrongdoer’s conduct. Fricker argues in favor of a possible solution to the concerns raised by private blame. She claims that “it is a straightforward feature of

---

<sup>61</sup> This is not even a mere hypothesis. Both Arpaly and Carlsson argue that the fundamental form of blame is private blame (Arpaly 2006; Carlsson 2017).

<sup>62</sup> See §5, Chapter One.



communicative acts in general—telling, warning, arguing etc.—that they can be withheld, kept private. [...] Non-communicated blame is therefore readily understood as derivative of Communicative Blame in just this simple way: sometimes it is better all things considered not to communicate a judgement even while it is of a type that is best understood as essentially apt for communication.” (Fricker 2016, 179). Admittedly, communicative acts such as telling or warning *can* be kept private, as Fricker argues. Moreover, I see no problems in conceding that sometimes we have prudential reasons not to outwardly express our warnings or other communicative acts. However, I fail to see how private blame can be “readily understood” as derivative of Communicative Blame, especially considering that the function of blame is – according to Fricker – to increase the alignment of moral understanding between the blamer and the person being blamed. It could be argued that private blame is derivative of Communicative Blame because it involves a reaction to wrongdoing that, *were it* outwardly expressed, it would discharge the main function of blame.<sup>63</sup> This solution, however, does not seem as straightforward as Fricker presents it. Moreover, it could be argued that Fricker offers *normative* reasons not to outwardly blame others. Clearly, these considerations play an important role in our decisions. However, I fail to see the role of these normative considerations in cases of blaming the dead. Fricker should also provide us with non-normative reasons in order to justify her claims about private blame being derivative of Communicative Blame.

Finally, dispassionate blame seems to be a problem for the parading-based account of blame. In her characterization of Communicative Blame, Fricker explicitly claims that emotions play an important role in increasing the alignment of moral understanding between the blamer and the person being blamed (Fricker 2016, 173). At the same time, she also argues that her paradigm-based theory of blame has a straightforward explanation of cases

---

<sup>63</sup> This is similar to the solution of cases of private blame offered by McKenna in his 2012.

of dispassionate blame: “the proper way to acknowledge the place of emotion is to locate it in our basic paradigm of blame, thereby crediting it as blame’s first impetus but without forcing us to conserve it artificially as a necessary condition of the mature and internally diversified practice of blame” (Fricker 2016, 171). Admittedly, it would be wrong to “artificially conserve” emotions as a necessary component of blame. However, it seems to me that emotions play a fundamental role in discharging the main function of blame. Recall what Fricker says about remorse (Fricker 2016, 173). In her characterization of Communicative Blame, it is emotions which discharge the role of blame – that is, increasing the alignment of moral understanding between the blamer and the person being blamed. In cases of dispassionate blame, *which element* should play the same role? It seems to me that the paradigm-based account of blame defended by Fricker relies too much on the role of emotions. Thus, it struggles in accounting for cases of dispassionate blame.

In this section, I have argued that the paradigm-based theory of blame defended by Fricker is unable to account for the varieties of our blaming practices. This account inherits the problematic aspects of the protesting account of blame, insofar as it struggles to account for private blame. Moreover, since in Communicative Blame emotions play an important role in explaining the function of blame, the paradigm-based account of blame faces difficulties in addressing cases of dispassionate blame. Ultimately, I think that the paradigm-based account of blame proposed by Fricker is to be rejected. In the next chapter, I defend a functional account of blame that provides a convincing explanation of blame and of our blaming practices.



## CHAPTER FOUR

### THE SIGNALING ACCOUNT OF BLAME

In the second and third chapters of this dissertation, I have argued that the theories of blame that I have taken into consideration cannot account for the main data points of blame. In other words, the problem of the nature of blame is still open. Is it possible to account for the varieties of blame, while still providing a satisfying theory of blame?

In this chapter, I defend a functional theory of blame that can explain the varieties of blame without giving away its explanatory power. The theory I defend in this chapter takes its cues from the functional account of blame offered by Shoemaker and Vargas in their 2021. They label this account the “Signaling Account of Blame” (Shoemaker and Vargas 2021). According to the signaling account of blame, the main function of blame is to signal the blamer’s normative commitments, which include a commitment to the enforcement of those commitments (Shoemaker and Vargas 2021, 582). This means that the signaling account of blame shifts the focus from the person being blamed and from the blaming interaction to the blaming agent. As I will argue throughout the chapter, this is an advantage of the signaling account of blame over other functional accounts of blame.

This chapter is structured in the following way. In §1, I present the theory of blame as signaling as a hypothesis. In the first section, my main aim is to motivate the choice of a signaling theory of blame by drawing an analogy with other practices studied by anthropology, ethnology, religious studies and even biology. In §2, I introduce the principles underpinning the Costly Signaling Theory. This theory offers me a promising framework to give a unifying explanation of blame and of our blaming practices. In §3, I focus on blame as a form of signaling, arguing that blame is a reliable signal of blamers’ normative competence and normative commitments. Moreover, I clarify the kind of benefits that blamers stand to gain by signaling their own normative competence and normative

commitments. In §4, I put to test the hypothesis of blame as signaling, evaluating its explanatory power against the data points of blame that I have identified in §5, Chapter One. Finally, in §5 I respond to objections that could be raised against the account of blame as signaling.

### **§1 The Hypothesis: Blame as a Signaling Device**

In this section, I present a hypothesis regarding the function of blame. This hypothesis has been suggested by Shoemaker and Vargas in a recent paper (Shoemaker and Vargas 2021). According to these authors, the function of blame is signaling the blamer's normative commitments.<sup>64</sup> Then, I describe various practices that can be explained by a single theory, namely the costly signaling theory. I conclude this section by motivating the idea that the same theory can be used to explain blame and our blaming practices.

In the previous two chapters of this dissertation, I have argued that recent theories of blame cannot account for the diversities of our blaming practices. To mention a few examples, emotional theories cannot account for dispassionate blame, while the relationship-based account defended by Scanlon seems unable to account for the problem of “blaming the stranger” (Wallace 2011; Sher 2013). Moreover, the Protest View held by Smith seems circular, insofar as it seems unable to offer a definition of protest that is independent from the concept of blame. These failures motivate us to look for an alternative account of blame. In order to introduce such an account, let me discuss a few cases drawn from the most disparate disciplines, such as anthropology, ethnology, religious studies and even biology. By the end of this section, I will argue that these cases can be explained by the same theory. Finally, my contention is that the same theory can account for the diversity of our blaming practices.

---

<sup>64</sup> “[Blame] is a costly response to norm violations defined most fundamentally not by any particular content—e.g., a mental state or activity—but by a function, namely, the signaling of the blamer's commitments, including a commitment to the enforcement of those commitments” (Shoemaker and Vargas 2021, 582).

Both in his 1975 and 2001, W. Irons discusses his own experience with the Yomut Turkmen of northern Iran. Irons claims that many religious practices and especially rituals act as signaling devices of commitment to norms, religious doctrine, and devotion to a group or community (Irons 1975; Irons 2001). Most of these practices and rituals, however, include particularly burdensome and costly activities. How can we explain the willingness of the members of a certain community to endure the costs associated with signaling practices? Irons argues that signaling leads to increased trust and cooperation among group members. This means that signalers stand to gain long-term benefits from their signaling activities. Thus, it can be argued that long-term benefits override the costs associated with certain practices. An example drawn from Irons's own observations may help us understand this point.

Irons lived two years and a half between the Yomut. In his writings about this experience, he focuses on the religious rites he assisted to. Irons frequently travelled with the Yomut, who every day had to stop five times in order to pray. Irons specifies that “the prescribed prayers had to be preceded with a ritual washing and had to be performed in a clean place while facing ‘The House of God’ in Mecca” (Irons 2001, 299). While traversing the desert, stopping five times a day and performing ritual washing can be particularly costly activities: “one had to interrupt travel, had to find water and a clean place for prayer, and usually had to make inquiries about the direction of Mecca” (Irons 2001, 299). Clearly, members of the Yomut community also pray while they are not travelling. However, Irons notes that he frequently had the impression that some members of the Yomut community were more regular in their prayers when they knew they had the possibility of being observed by others. These highly regulated praying practices create strong bonds between members of the same community, who must stop all their activities in order to gather and pray.

The five-times-a-day-prayer is not the only signal that Irons describes in his works. He also notes that Yomut are particularly strict when it comes to Ramadan, “the month of fasting, when all adults (except pregnant or lactating women or those who were ill) fasted from sunrise to sunset” (Irons 2001, 300). This practice is particularly costly. During the month of fasting men and women have to wake up before sunrise in order to consume a large meal, and adults have to work without having the possibility to eat throughout the day. However, this prolonged period of fasting also has its long-term benefits: members of the Yomut community strengthen their social bonds by meeting after sunset in order to consume large meals together. Finally, Irons describes another costly practice the Yomut engage in: “all Muslims are required annually to give a defined portion of their wealth to the poor or to those who devoted themselves full time to the study or teaching of religion” (Irons 2001, 301). All these practices not only reinforce the religious commitment among members of the Yomut community, but they also strengthen commitment to a set of rules concerning appropriate moral behavior.

Of course, members of the Yomut community are not the only human beings who regularly engage in costly practices in order to signal their commitment to rules, norms, and principles. Shoemaker and Vargas describe the phenomenon of torch fishing (Shoemaker and Vargas 2021, 581-582). Fishermen of Ifaluk – a coral atoll in the Pacific Ocean - spend a lot of time and energy in order to engage in torch fishing, which consists in a highly ritualized form of fishing that takes place at night. Fishermen who engage in torch fishing at night do not suspend their regular daytime fishing activities. This means that fishermen who engage in torch fishing almost do not sleep at night. Moreover, it has been observed that during the night fishermen manage to gather way less fish than the fish gathered during daytime fishing (Sosis 2001). Why do some men from Ifaluk engage in torch fishing, provided that this activity proves to be very costly, and involving no immediate payoff for fishermen?

Again, the explanation lies in the role of signaling. Fishermen who engage in the activity of torch fishing signal their personal qualities that are otherwise difficult to exhibit to other members of their community. In participating to torch fishing, men from Ifaluk convey precious information about their physical qualities, their ability to resist stressful conditions, and their capabilities to face difficult challenges. Conveying this information helps them gain a higher social status, and the opportunity of getting into beneficial marriages (Sosis 2001). Torch fishing is a costly and burdensome practice, but those who engage in it stand to gain long-term benefits by signaling their qualities to observers and other members of the community.

Interestingly, human agents are not the only living beings who perform costly activities in order to send signals. We are all familiar with the phenomenon of autumnal leaf color. Some biologists observed that trees invest a lot of energy in order to turn their leaves into the typical autumn colors - mostly brown, red, and yellow (Archetti 2000; Hamilton and Brown 2001). This could be explained by considering these colors as signals of powerful chemical defenses. According to Archetti, Hamilton and Brown, trees signal to aphids – a species of pests – that they are well-equipped against parasites. In turn, this may help trees reduce their parasitic load.

In this section, I have shown that a vast number of practices can be explained by referring to the notion of signal and the related act of signaling. As I have argued in the opening of this chapter, my aim is to explain blame by referring to the same notion. Our blaming interactions have many elements in common with the practices I have been describing so far. Oftentimes – even if not always – blamers find themselves in unpleasant emotional mental states. Moreover, blame may involve – as Scanlon suggests – a modification of an ongoing relationship (Scanlon 2008). As Hieronymi describes in her 2004, blame has a characteristic force or “sting” (Hieronymi 2004). All these elements described



by different content-based accounts of blame can be defined as costs associated with blame. Most content-based theories of blame try to capture and explain these phenomena, but they fail in accounting for the varieties of blame. A signaling account of blame could easily account for these phenomena. According to the signaling theory of blame, all these features of blame can be described as costly. Moreover, blame can be seen as a signaling device blamers use in order to convey information about their own personal characteristics, such as their commitment to norms that they have internalized and that they are ready to enforce against transgressors. Finally, blamers can stand to gain long-term benefits from their blaming practices.

In §3, I will clarify why I think it is plausible to apply the framework provided by the costly signaling theory to blame and to our blaming practices. Before doing so, however, I need to discuss the details of the costly signaling theory. Up until this point, I have only drawn a vague analogy between our blaming practices and other activities that can be explained by the costly signaling theory. In order to make this analogy clearer – and thus build a signaling theory of blame – it is necessary to provide the detailed theoretical framework of the costly signaling theory. In the next section, I will present such a framework.

## **§2 What does Signaling Mean? The Costly Signaling Theory**

The Costly Signaling Theory – from now on, CST – was formally developed by the economist M. Spence.<sup>65</sup> Spence applied CST to behavioral economics, in his attempt to explain how individuals who have asymmetrical information share and communicate between them in a given market (Spence 1973; Spence 2002). In the previous section, I have shown that CST proved to be an efficient framework when developed and applied in different areas, such as anthropology (Irons 1975; Irons 2001), evolutionary psychology

---

<sup>65</sup> Early discussions of signaling theories can be found in Veblen 1899 and Mauss 1924.

(Bliege Bird, Smith and Bird 2001), sociology (Eriksson 2019), and religious studies (Irons 2001). My contention is that the philosophical research on blame can benefit from the framework provided by CST. Thus, I will now present a detailed account of this theory.

In both his 1973 and his 2002, Spence describes a widespread phenomenon. In most markets, individuals have asymmetrical information. This means that a group of individuals possess certain information that is unknown to other individuals (Spence 1973). Sometimes, individuals can benefit from keeping this information for themselves, and thus they avoid sharing it. Under certain circumstances, however, individuals may benefit from sharing this information. Job candidates, for example, would benefit from letting their employers know their capabilities and their skills. By sharing relevant information about themselves, prospective employees can maximize their probability of being hired. Some information, however, may be difficult to convey. This is where CST comes into play. Spence argues that candidates can *signal* precious information about themselves to the employers, especially when merely *communicating* the very same information would not be effective (Spence 1973, 358).

Let us apply CST as developed by Spence to a case we are all familiar with. The academic job market is notoriously competitive, and young researchers may greatly benefit from signaling their skills and capabilities to hiring committees. When applying for a research position, postgraduates send their CV in order to signal to the hiring committee that they possess relevant skills. The University from which they have obtained their PhD may be an important signal, as well as their list of publications. Spence himself observes that education is an important signal of the qualities of potential employees; qualities that may be difficult to assess otherwise. Spence also draws the distinction between indices and signals. Indices are “observable, unalterable attributes”, while signals are defined as “observable characteristics attached to the individual that are subject to manipulation by him” (Spence

1973, 357).<sup>66</sup> Despite being at least in principle observable, certain characteristics are difficult to assess. Universities' hiring committees certainly cannot assess a researcher's skill by simply looking at him or her. However, committees can evaluate candidates' CVs, which are instruments through which candidates signal their skills and capabilities. Receiving a PhD from a prestigious University is a strong signal of the capabilities of a candidate. The same can be said of publishing in a highly ranked journal in a specific field. If they want to raise their possibilities of being hired, candidates are strongly encouraged to publish high quality papers in impactful journals. However, things are not this simple. We are all familiar with the difficulties linked to publishing a paper on highly ranked journals. In other words, publishing papers on prestigious journals is a costly activity. Nonetheless, researchers can stand to gain long-term benefits from being able to send such strong signals, such as increasing their possibilities of being hired. Moreover, these signals are reliable sources of information about the set of skills possessed by candidates. In order to publish their works in highly ranked journals, early-career researchers have to develop compelling arguments. Additionally, their writing must be concise and clear. These are valuable skills that are costly to acquire, as they require a lot of time and energy. An outstanding list of publications is a reliable signal of the qualities of researchers precisely because it is hard to fake or to mimic.<sup>67</sup> This points to an interesting distinction recognized by proponents of CST. Signals can be honest or dishonest. An honest signal is a reliable source of information for the receiver. A signal can be honest whether because it is impossible to fake, or because the costs associated with faking it exceeds the benefits (as in being caught cheating about your publications). Dishonest signals, on the other hand, are not sources of reliable information (Dawkins and Krebs 1978). By honestly signaling to the hiring committee their capabilities and their skills, candidates raise their

---

<sup>66</sup> The candidate's height or eye color, thus, can be classified as indices. Attributes like age can change (and thus are not unalterable), but they are nonetheless indices because they do not change at the discretion of the individual (Spence 1973, 358).

<sup>67</sup> Of course, one could blatantly lie about his or her list of publication. This, however, is hardly ever a good strategy if one aims at being a professional researcher.

probability of being successfully hired. Costly signaling theory explains the reliability of a signal by referring to the degree of difficulty of faking that signal: the more a signal is hard to fake, the more it is a reliable source of information about the signaler's own characteristics.

In the previous section, I have taken into account three different phenomena that can be explained by CST. Despite being diverse in many aspects, the religious rituals of the Yomut, the torch fishing of Ifaluk inhabitants, and even autumn leaf color can be explained by referring to the costly signaling theory. In this section, I have presented CST as it has been developed by Spence; moreover, I have applied CST to a familiar phenomenon in the life of every academic. Now, it is time to generalize from these examples and pin down the characteristics of CST. The theoretical framework of costly signaling theory can be applied when:

- Members of a given community have qualities that others find hard to perceive.
- In order to convey important information about themselves, agents may recur to signaling practices.
- Signals are perceived as costly, e.g., they involve efforts on behalf of the signalers.
- Despite the costs associated with signals, signalers stand to gain long-term benefits from conveying information about their own personal characteristics.
- Signals are hard to fake, and thus reliable. In turn, the costs associated with signaling discourage free riders from taking advantage of the practice.

These five characteristics are summed up by Shoemaker and Vargas, who write that “when a costly signal becomes part of some stable system, it will be one in which it has observers, it is hard to fake (otherwise it would be too easily imitated), it delivers accurate information to the observers, and it benefits the signaler” (Shoemaker and Vargas 2021, 586).

In the introductory section to this chapter, I claimed that my aim is to identify the function of blame with a form of signaling. In this section, I have provided the details of CST. The next step will be verifying whether blame and our blaming practices actually present characteristics that can be convincingly explained by the costly signaling theory. In the next section, I turn to this task.

### §3 Blame as Signaling

The costly signaling theory can be used to explain a wide range of phenomena. Following Shoemaker and Vargas, my contention is that the framework provided by CST can also be applied to the study of the function of blame. In this section, I take a close look at our blaming practices in order to verify if the framework provided by CST can be applied to them. By the end of this section, it should be clear that CST provides a powerful tool in order to give a unifying explanation of blame and of our blaming practices. The main function of blame is to signal the blamer's normative competence and her normative commitments. Through blame, blamers convey honest information about their own normative commitments, along with their readiness to enforce these commitments (Shoemaker and Vargas 2021, 582). In order to defend this claim, I will answer the following questions: what information about themselves do blamers convey in blaming each other? How is blame costly? What long-term benefits do blamers stand to gain? What makes blame hard to fake, and thus a reliable source of information about the blamers?

Let me start by answering the first question: what do blamers signal? I have already anticipated that the function of blame is signaling the blamer's normative competence and normative commitments, including the commitment to enforce these commitments. Let me now clarify this claim. Most philosophers see blame as a reaction to something of negative normative significance about someone or their behavior (Coates and Tognazzini 2018, 1).

Let us imagine that agent B violates a norm, and A reacts to this norm violation by blaming B. In blaming B, A signals that she is normatively competent in relation to the normative domain that has been violated (Shoemaker and Vargas 2021, 589). Agents who are normatively competent in a given normative domain are often defined as agents possessing the ability to respond to normatively relevant features of their environment (Talbert 2012, 89-90).<sup>68</sup> By blaming someone, blamers signal their normative competence in the relevant normative domain. Being normatively competent in a given domain is an information about oneself that would be difficult to convey otherwise. Through blame, however, we are able to signal our normative competence in a given normative domain. This is not the only important information that blamers can signal through blame and by engaging in blaming interactions. Blamers also convey crucial information about their own practical identity, about the things they value, about their loyalty to certain groups and community and, finally, about the weight they give to norms violations. It should be added that blamers signal this information about themselves both voluntarily and involuntarily. Blamer may feel emotionally engaged by certain violations of norms, thus reacting with anger or strong indignation. These emotional responses are typically hard to control, but they nonetheless send a lot of signals both to the person being blamed and to bystanders.<sup>69</sup> Finally, blamers do not only show their normative competence. As Shoemaker and Vargas correctly underline, blamers also signal their commitment to norms; in turn, this commitment is accompanied by the blamers' willingness to policy norms violations and the willingness to enforce the norms (Shoemaker and Vargas 2021, 589-591). As I will clarify below, blame does not only signal that the blamer *knows* about the existence of a norm; blame also signals that the blamer has internalized the relevant

---

<sup>68</sup> The concept of “normative competence” is often at the center of discussions about moral responsibility. Because of this, most authors refer to the concept of “moral competence” (Wallace 1996; Watson 2004a; Wolf 1990; Wolf 2003). I prefer using the term “normative competence”, in order not to restrict the scope of my analysis to moral blame.

<sup>69</sup> The involuntary aspects of signaling are of particular importance in considering the reliability of signals. These involuntary aspects are typically hard to fake, and thus they contribute to the reliability of signals.

norm. After having invested time and energy in the interiorization process, blamers are committed to the set of relevant norms. Blame, as I have argued, is a reliable signal of this normative commitment.

Now, I want to address the second question: how is blame costly? What are the costs associated with blame? My contention is that blame can be costly in at least two distinct ways. First, blaming interactions can be costly both for the blamer and the person being blamed. Second, in order to send honest signals, blamers have to internalize the norms they are committed to, and the internalization process can be costly. Thus, blame can be costly even outside of blaming interactions. Let us start by considering the costs associated with blaming interactions. Most authors note that being on the receiving end of blaming activities can be an unpleasant experience, while some even contend that being blamed is harmful (McKenna 2012, 134-141; Matheson and Milam 2021). However, here I want to strengthen my position by focusing on the costs associated with being the blamer. Albeit blame does not always involve negative emotions such as resentment and indignation (as most emotional accounts of blame maintain), blamers who engage in blaming interactions often feel these emotions (McGeer 2013).<sup>70</sup> Albeit not being necessarily hurtful, feeling negatively valued emotions such as indignation or resentment can be described at least as costly.<sup>71</sup> Additionally, blame is not only costly because it typically involves emotional investment on behalf of the blamer. Blame is also costly because it can be disadvantageous for the blamer to engage in an overt blaming interaction. For example, there can be prudential reasons to refrain from blaming others.<sup>72</sup> Blame can negatively affect personal relationships, as it can lead to relationship

---

<sup>70</sup> This is an additional advantage of the signaling account of blame over content-based accounts of blame. The signaling account of blame can account for instances of dispassionate blame while still being able not to detach itself from human psychology. For a discussion about the psychological aspects of blame, see McGeer 2013.

<sup>71</sup> It is not even necessary to define feeling resentment or indignation as intrinsically costly. Feeling such emotions can be burdensome even in an instrumental way: consider the psychological distress caused by feeling hostile or strong emotions.

<sup>72</sup> Note that this can still be true even if the person being blamed is blameworthy. This is an additional reason to elaborate a theory of blame that can convincingly explain the distinction between judgments of blameworthiness and active blame.

impairments or even the interruption of a personal relationship. Moreover, blame can lead to potentially unfavorable conditions for the blamer – such as when, for example, an employee overtly blames her boss. The costs of blame, however, are not limited to blaming interactions. In order to become the kinds of agents who are able to send honest signals, blamers had to invest enormous social and psychological resources (Shoemaker and Vargas 2021, 586). When I blame an agent, I am doing more than merely noting that a norm has been violated. Merely noting norms transgressions would require nothing more than the memorization of said norms. Blaming, on the other hand, requires more than memorizing norms and registering transgressions: genuine blame involves a process of internalization of such norms.<sup>73</sup> This process is not only costly *per se*, but blamers also have to invest social and psychological resources to keep up with “what the incredibly wide-ranging and subtle interpersonal norms for interpersonal interaction are, for they regularly evolve” (Shoemaker and Vargas 2021, 586). Moreover, blamers who are honestly normatively committed are also typically prone to enforce the norms violated by the person being blamed; this represents an additional cost. Thus, we can conclude that blame involves great costs on behalf of blamers.

Let me move to the third question. In presenting CST, I have noted that signalers stand to gain long-term benefits from their signaling activities. How can this be true in the case of blame? Or, in other words, how can blame benefit blamers in the long run? Recall that the function of blame is to signal the blamer’s normative competence and her normative commitments. By signaling their own normative commitments, blamers can build a trustworthy reputation. By blaming a liar, for example, the blamer signals that she is the kind of person who cannot stand lies and liars. In turn, this can help the blamer build a trustworthy

---

<sup>73</sup> This could also help us distinguishing blame from criticism without having to revert to emotional accounts of blame. While criticizing, I am pointing out to the wrongdoer that different courses of action were available, and that some of these courses of action were preferable. Blame, on the other hand, requires a form of engagement on behalf of the blamer. This form of engagement can be explained as having endured a process of internalization of the norms being violated by the wrongdoer.



reputation. As Shoemaker and Vargas also note, being known as a trustworthy person does – at the very least – provide a solution to prisoner’s-dilemma-type situation the blamer may find herself in (Shoemaker and Vargas 2021, 587). However, blame is more than a mere instrument of reputation management. Human beings typically care about being recognized as normatively competent creatures. As I have argued above, blame provides an excellent device for blamers to convey precious information about themselves, such as their normative commitments and their membership in a specific community.<sup>74</sup> And, as I have already noted, information about these personal characteristics is difficult to convey otherwise. Despite the costs associated with blame, being recognized as a normatively competent agent is something that few would renounce to. Blamers who pay the costs associated with blame also manage to be part of normative communities. To illustrate this point, let us focus on moral blame. Blamers who have internalized moral norms and are disposed to policy transgressions of moral norms are part of the moral community. Being part of the moral community vastly overcomes the costs associated with blame and with the process of internalization of moral norms. Moreover, the alternative – that is, being excluded from the moral community – is even more costly. Being excluded from the moral community is something very few would accept. In Strawsonian terms, accepting not to be part of the moral community would mean accepting that others adopt an objective attitude or stance towards us (Strawson 2008, 9). In order to take part in meaningful human relationships, agents have to be normatively competent. And, as should be clear by now, blame conveys important information about our normative commitments and about our normative competence. Thus, blamers stand to gain long-term benefits from their signaling activities, despite the costs associated with blame.

Finally, let me explain why blame is a reliable source of information about the blamer. Just like an outstanding list of publications is a reliable signal of the skill and capabilities of a

---

<sup>74</sup> In blaming, blamers also convey precious information about how they expect others to treat them.

prospective researcher, blame is a reliable source of information about the normative competence and the normative commitments of the blamer. Why? The explanation follows from my analysis of the cost associated with blame. As I have already argued, blame is costly both for the blamer and the person being blamed. Moreover, the process of internalizing the norms is highly costly. These combined costs ensure that the signals issued by blamers are hard to fake and thus reliable. This also disincentivizes free riders to exploit our blaming practices in order to gain long-term benefits. Imagine a situation in which agent “A” fakes her normative commitment to a norm “N” that has just been violated by agent “B”. A blames B for this violation, despite not being committed to N. If she does so overtly, B or some bystander can easily detect A's attempt at faking her normative commitment to N. Faking such blaming interaction would require A to correctly mimic all the attitudes, facial expressions, bodily movements, and even voice intonations that are frequently associated with blame. Not only fake blame is easily detected by observers, but faking blame could result in extremely negative consequences for the faker. Think of your reaction at discovering that your spouse has been repeatedly faking her blame towards persons who cheat on their partner. As Fraser aptly notes, the reliability of signals can be maintained not only because signals are costly *per se*, but also because dishonesty – while possible – is not profitable (Fraser 2012). These considerations help us understand why blame is a reliable source of information about blamers’ normative commitments and normative competence. Blame is a reliable signal of the blamer’s normative competence and normative commitments because it is hard to fake in a convincing way; moreover, it is also unprofitable for free riders to fake blame. Thus, agents are disincentivized to mimic honest blame, rendering blame an even more reliable source of information about blamers’ normative competence and normative commitments.

In this section, I have enquired whether our blaming practices can be explained by referring to the framework provided by costly signaling theory. By addressing four important questions, I have shown that the framework provided by CST can be applied to blame and to our blaming practices. Thus, we can conclude that the varieties of blame can find a unifying explanation in a functional account of blame that stresses the signaling function of blame. The main function of blame is to signal the blamer's normative competence and normative commitments. Now that I have shown that blame can be explained by referring to CST, I want to strengthen my argument by testing the theory of blame as a form of signaling.

#### **§4 The Data Points of Blame and Blame as Signaling**

In this section, I put to test the signaling theory of blame against the data points of blame that I have identified in §5, Chapter One. Is the signaling account of blame that I have defended throughout this chapter able to account for the data points of blame?

Let me start by considering the claim according to which blame consists in more than the mere recognition that the person being blamed has acted in the wrong way. I think that the signaling account of blame can easily account for this data point. I have claimed that when we blame someone, we signal our normative commitments while also signaling that we are normatively competent agents. At the same time, we signal our willingness to enforce the norms we are committed to. This element of self-disclosure is precisely what distinguishes blame as signaling from the mere recognition that an agent has acted in the wrong way. As I have argued in the previous section, in order to merely register an agent's faulty behavior, it is not necessary to convey information about our own normative commitments. Even individuals who are not normatively competent can register the violation of a norm. Blame, however, is different, for it involves an element of self-disclosure and personal involvement

from the blamer. And the signaling account of blame can easily account for this element, since it explains blame as an honest signal of the blamer's normative commitments.

Following the order of data points that I have presented in §5, Chapter One, I turn now to directed dyadic blame. In this blaming interaction, the blamer directly addresses the person being blamed. Moreover, she does so by overtly manifesting her blame. The paradigm case is this: I wrong you, and you blame me for my wrongdoing. Contemporary accounts of blame rarely struggle to account for directed dyadic blame: some accounts even consider it to be paradigmatic or explanatorily prior to other forms of blame (McKenna 2012; Fricker 2016). In this respect, I do not expect the signaling account of blame to perform better than other accounts of blame. Quite simply, in cases of directed dyadic blame the blamer signals her normative commitment directly to the person being blamed. If agent B wrongs A, A could react by blaming B. Under some circumstances, this means that A angrily reproaches B and demands for an apology. Or, as Scanlon would say, A modifies her relationship with B. These elements, however, are not necessary components of blame. According to the signaling account of blame, all that is needed in order to properly speak of blame is the fact that A addresses B by signaling her commitment to the norm(s) violated by B. Thus, the signaling account of blame can account for directed dyadic blame.

Third, I want to discuss third-party overt blame. We should keep in mind that when we blame, we do not always directly address the wrongdoer. We do not directly address the wrongdoer when the blamer and the person being blamed are physically distant, or when the blamer has prudential reasons not to directly address the wrongdoer. Most contemporary accounts of blame focus on directed blame, losing part of their explanatory power when they have to address cases of third-party blame. The signaling account of blame, on the other hand, does not need to give away its explanatory power in order to explain cases of third-party blame. Whether we are directly addressing the wrongdoer, or blaming an agent while

interacting with others, we are still signaling our normative competence and our normative commitments. Moreover, the signaling account of blame also helps us understand how the same blaming interaction may communicate different things to different agents. Sometimes, blamers blame an agent by interacting with more than one person. Different members of the audience could pick up different information from this blaming interaction. The signaling account of blame can easily account for this phenomenon: different members of the audience pick up different signals, even if the blaming interaction is the same. Thus, the signaling account of blame does not give away its explanatory power when it shifts from directed dyadic blame to third-party blame. On the contrary, this functional account of blame can explain the nuances of our blaming interactions.

Now, let me move to private blame. As Shoemaker and Vargas also note, it can be thought that private blame might constitute a problem for a signaling account of blame (Shoemaker and Vargas 2021, 593). The signaling account of blame states that the main function of blame is to signal the blamer's normative commitments, along with the blamer's disposition to enforce violated norms. It may be objected that, in private blame, the blamer does not signal her normative commitments to anyone. In private blame, the blamer does not express her blame: how can she signal anything if she does not outwardly express her blame? There are at least two answers to this question. The first answer is to deny that, in private blame, the blamer does not send signals to anyone. Even in private blame, the blamer signals her commitment to violated norms to at least one individual: herself. This may reinforce her own commitment to the violated norms. I will discuss this point at a greater length in the following paragraph, where I focus on self-blame. The second, more sophisticated answer, is focusing on the fact that, in private blame, the blamer is still sending signals even if there is no one picking up these signals (Shoemaker and Vargas 2021, 594). While a private activity of communicating is hardly intelligible (Macnamara 2015b: 217;

Chislenko 2019), a private activity of signaling has a straightforward explanation. This has to do with the process of norms internalization, which in turn constitutes a fundamental component of blame as signaling. When we internalize a norm, we are disposed to react to violations of that norm by blaming the violator: “if one is truly committed to the norms in question, then one will be disposed to produce a blame signal in response to their violation in all sort of circumstances, even where there is no external audience to pick up this signal” (Shoemaker and Vargas 2021, 594). In the following section, I will come back to the problem of private blame. Now, let us move to discuss the next data point.

Now, we could formulate the following question: how can a signaling theory of blame account for cases of self-blame? It could be objected that, in cases of self-blame, the blamer does not signal her normative commitments to anyone. After all, cases of self-blame are describable as cases in which the blamer and the person being blamed coincide. Thus, it may be argued that in cases of self-blame the blamer only signals to herself. But this is not necessarily true. In truth, I think this objection misses an important feature of the signaling account of blame. Even if the blamer and the person being blamed are the same individual, it does not mean that observers cannot pick up signals emitted by the blamer. If I outwardly blame myself for some faults of mine, I convey a lot of information about my own normative commitments. Others would perceive that I am strongly committed to the norms I have violated; so much committed, in fact, that I am disposed to outwardly blame myself.<sup>75</sup> However, there can be cases of private self-blame, where not only the blamer and the person being blamed coincide, but blame is not being outwardly expressed. These cases may present a challenge to the signaling account of blame. For this reason, I will return to this issue in the next section.

---

<sup>75</sup> In this spirit, it could be argued that self-blame which is outwardly expressed is particularly costly.

We can now turn to the problem of blaming the distant. In Chapter One, I have argued that we can distinguish two cases of blaming the distant. First, we can blame agents who are distant in space – that is, agents who are not physically present during the blaming interaction. Second, we can blame agents who are distant in time, e.g., the dead. The first case does not pose a particular challenge to the signaling account of blame. According to this account, it is not necessary that the person being blamed picks up the signal. Others can pick up the signal sent by the blamer. Imagine meeting your coworkers at the pub after work. While talking, one of your coworkers blames your boss for her insensitive behavior. Even if your boss is absent, your coworker is still signaling his own commitment to the norm(s) violated by your boss. He is making it clear that he stands up for certain values, and that he does not tolerate the violations of certain norms. Evidently, he is not signaling all this to your boss: he is sending signals to his own coworkers, including you. The signaling account of blame can also easily accommodate for the phenomenon of blaming the dead. Communicative accounts of blame typically face problems when they attempt to explain this phenomenon. In cases of blaming the dead, the blamer clearly does not issue a demand for apology to the person being blamed. At the same time, it would also be odd to claim that we blame the dead “to bring increased alignment of the moral understandings of wronged and wrongdoer”, as Fricker claims (Fricker 2016, 174). The signaling account of blame has a straightforward solution to the problem of blaming the dead. When we blame historical figures – or, more generally, the dead – we signal our normative commitments to the living. In blaming a despotic tyrant from a distant past, we signal our commitment to values such as freedom and equality. When we blame historical figures from the past, we make it clear that we stand up for certain values and that we are ready to enforce the norms violated by the agent being blamed. Clearly, we do not convey this information to the agent being blamed, since she is dead long ago. However, the living may pick up these signals and gain a better understanding of our normative commitments. In §5, Chapter One, I have raised an

issue regarding the phenomenon of blaming the dead. It may be argued that we can properly speak of blame only in the cases of recently deceased agents, especially those whose actions still have negative consequences in the present. On the contrary, it could be argued that we do not properly blame violent Roman emperors or ancient tyrants. We may disapprove of their violent conduct, or even criticize their behavior. These reactions, however, do not amount to blame. This argument is usually tied to cases of dispassionate blame. Emotional theories of blame, as I have shown in the second chapter of this dissertation, claim that blame identifies with an emotional mental state. Oftentimes, however, we do not feel any blaming emotion towards historical figures who died thousands of years ago. Thus, according to emotional theories of blame, we cannot blame agents who died long ago. We may judge these agents to be blameworthy, and we may disapprove of their deeds. However, I think that it is an error to exclude the possibility of blaming ancient historical figures. The signaling account of blame is well equipped to account for this possibility. According to this account of blame, when we blame an agent – being it alive, recently deceased, or dead for a thousand years – we signal our commitment to norms violated by that agent. This signal, in turn, can be picked up by the living. Even if Spanish Conquistadores died five hundred years ago, we can blame them by standing up for values such as freedom and equality, as well as signaling to others our commitment to norms violated by the European colonialists. Thus, I believe that the signaling theory convincingly accounts for the varieties of blame without giving away its explanatory power.

Finally, let me discuss cases of dispassionate blame. The theory of blame as signaling has a ready explanation for cases of dispassionate blame. When we blame, it is not necessary that we feel certain emotions such as resentment or indignation. What distinguishes blaming interactions from other types of reactions to wrongdoings is the fact that, while blaming, the blamer signals her normative competence and her normative commitments. In order to send



signals, emotions are not necessary. Undoubtedly, many blaming interactions include certain kinds of emotional responses. Emotions at least partly constitute the costs associated with blame and are also part of what makes the signal hard to fake. But even so, blamers do not necessarily need to feel emotions such as resentment or indignation in order to send signals about their normative commitments. By not including emotions among the necessary components of blame, the signaling account of blame can account for cases of so-called dispassionate blame. At the same time, the signaling account of blame does not need to depart from human psychology by maintaining that emotions do not play any part in our blaming activities. According to the signaling account of blame, emotions still play an important role as commitment devices (Frank 1988, 4-7).

In this section, I have argued that the signaling account of blame can account for the data points of blame that I have identified in §5, Chapter One. In my opinion, the signaling account of blame has a straightforward explanation for every data point on the list. Compared to the other accounts of blame that I have discussed throughout this dissertation, the signaling account of blame has a greater explanatory power. At the same time, this account explains in a unifying way the diversities of our blaming practices. In the next section, I present some possible objections to the signaling theory of blame.

### **§5 Objections to Blame as Signaling**

In this section, I want to address some possible objections to the functional theory of blame that I have defended in this chapter.

In the previous section, I claimed that the signaling account of blame can also account for the phenomenon of private blame, e.g., blame that is not outwardly expressed. But how can blamers signal something – namely, their normative commitments and their

normative competence – if they do not even outwardly express their blame?<sup>76</sup> In other words, this concern about private blame could be expressed in this way: *to whom* is the blamer signaling her normative commitments, if she does not outwardly express her blame? In the previous section, I argued that the signaling account of blame is well-equipped to answer this question. First, note that most functional accounts of blame allow for isolated instances of blaming interactions in which blame does not discharge its main function (Fricker 2016; Sliwa 2021). According to functional theories, a hammer is still a hammer even if it is not being used to drive a nail in the wall.<sup>77</sup> However, proponents of the signaling account of blame could not explain *every* instance of private blame as an exception to the main function of blame. Fortunately, there are more convincing answers to the concern I raised about private blame. While we privately blame an agent, it is simply not true that we do not send signals to anyone. Even if we do not express our blame, it is still true that we signal our normative commitments to ourselves. These signals are still costly, for we have internalized the norms that have been violated by the person being blamed. At the same time, blamers who privately blame can still stand to gain long-term benefits. By being the kind of agents who are *disposed to blame* ourselves and others – even if privately – we pay the costs associated with being part of the same community. And, as I have noted in §3, this sense of community is considered to be a great benefit for human beings. Thus, even instances of private blame retain the characteristics of blame as signaling that I have described throughout this chapter. Conversely, other communicative accounts of blame – such as those I have analyzed in the previous chapter – seem unable to account for instances of private blame. In particular, they

---

<sup>76</sup> Note that this concern could also be raised in cases of private self-blame. Instances of expressed self-blame, however, are easily addressed by the functional account of blame that I defend in this chapter. Imagine a chess player losing her temper after realizing her blunder. She could take her head between her hands and whisper to herself “how could you not have thought about *that* move?!” In this scenario, the blamer and the person being blamed coincide. However, the chess player is not only signaling something to herself. She is also conveying important information to observers, who may include both her opponent and bystanders.

<sup>77</sup> For reference, see §1, Chapter Three, in which I present the notion of “communicative entity” that does not necessarily discharge its function (Macnamara 2015b).

cannot explain *why* we should *blame* someone, provided that in cases of unexpressed blame the only audience of the blaming activity is the blamer itself. What distinguishes an instance of communication from an instance of blame, if the only function of blame is to communicate something? The signaling account of blame, on the other hand, has the resources to explain the difference between merely communicating and signaling. As I have argued, blamers have had to invest significant psychological and social resources in order to be the kind of agents who are able to send signals about themselves – that is, they had to internalize norms. In turn, this process of internalization of norms is what distinguishes merely communicating from signaling. Thus, we can conclude that the signaling account of blame – despite seeing the main function of blame as communicating something – is better equipped than other functional theories of blame to account for cases of private blame.

Now that concerns about private blame have been addressed, I can move to other possible objections to the theory of blame as signaling. Another concern about the signaling account of blame that I have defended in this chapter is the following. I have argued that emotions play an important role in the signaling account of blame. Namely, emotions are strong commitment devices (Frank 1988, 4-7). The objection that could be raised is the following: why don't we drop the signaling function, and maintain a sophisticated version of an emotional account of blame that highlights the role of emotions as commitment devices? The answer to this concern is fairly simple. In Chapter Two of this dissertation, I have argued that content-based accounts of blame share an important issue – namely, the inability to give a unifying explanation of our diverse blaming practices. Emotional accounts of blame – even in their most refined versions – still have to satisfactorily explain why our intuitions about cases of dispassionate blame are wrong. As I have already argued, emotional accounts of blame rule out the possibility of dispassionate blame. However, it seems plausible to think that we may blame each other even without the need to feel emotions such as resentment,

indignation, and guilt. More importantly, it has been argued by many that a convincing theory of blame has to account for cases of dispassionate blame (Scanlon 2008; Sher 2006; Smith 2013). Thus, I see no reason to abandon the signaling account of blame in favor of a refined emotional account of blame that puts at the center the role of emotions as commitment devices. On the other hand, I still think that emotions typically play an important role in our blaming activities. As I have argued in §3, emotions are part of what makes blame a reliable signal of the blamer's normative commitments. For this reason, I think that the signaling account of blame that I have defended in this chapter does better than other functional accounts in taking into consideration the psychological features of blame (McGeer 2013). At the same time, the signaling account of blame does not commit itself into maintaining that there are no instances of dispassionate blame.

Third, I want to consider a more substantial objection to the theory of blame as signaling that I have defended in this chapter. According to this objection, there seems to be no particular reason to recur to *blame* in order to *signal* our normative commitments. There seem to be many different ways in which agents can signal their normative commitments. In other words, why should we see *blame* as a device to convey precious information about ourselves? As I have showed in §1, many human interactions can be interpreted as signaling activities; why should be blame explained as a signaling device, too? Why don't we forego blame in favor of other means to signal our normative commitments?<sup>78</sup> This is an important objection, but I think it ultimately underestimates the explanatory power of the costly signaling theory of blame. As I have shown in §2, CST does not generically state that agents convey precious information about themselves by communicating with others. According to CST, signalers convey information about themselves *that would be difficult to convey otherwise*. If

---

<sup>78</sup> For a discussion about the reasons we may have to keep blaming each other, see Franklin 2013. Franklin attempts to defend an even more ambitious thesis, according to which blame is not only justifiable or permissible, but even good or right.

we were to forego blame in order to convey information about our normative commitments in some other way, where should we look at? Of course, merely announcing that one is normatively competent and normatively committed would not be sufficient, for it would be too easy for free riders to exploit this practice and fake their own agential qualities. Merely noting or acknowledging that an agent has violated a norm does not fare any better, for this would not be enough to signal one's normative commitments. An agent could easily try to fake her normative commitments by repeatedly noting other agents' transgressions, for merely registering agents' faults does not require a process of interiorization of norms. But, as I have argued in §2, CST specifies that signals are reliable sources of information because they are hard to fake. After all, we should not forget that signals are costly. Thus, to merely register other agents' faults would not be a reliable signal of one's normative commitments.<sup>79</sup> Finally, it could be argued that emotional reactions to wrongdoing such as horror, anger, or disappointed could be honest signals of one own's normative commitments. After all, certain emotional reactions are undoubtedly costly. Moreover, it takes time to internalize norms so as to react with strong emotions such as horror or anger to transgressions of relevant norms. Emotional reactions, additionally, would be very hard to fake. The reason they are so hard to fake is that they are hard to control in the first place. According to CST, however, signals have to be consciously controllable in order to be reliable sources of information about signalers' characteristics (including signalers' normative competence and normative commitments).<sup>80</sup> Sincere emotional reactions would be correctly classified as indices, or signs of one's normative commitments. Thus, the framework provided by CST does not seem applicable to emotional reactions such as horror and anger. As I have been arguing throughout the chapter, blame seems to be a perfect fit for CST. The function of blame is

---

<sup>79</sup> It could also be argued that registering other agents' faults would not be a reliable signal of one own's normative competence. Psychopaths, for example, are considered to be able to register faulty actions performed by others, even though they lack normative competence (Fine and Kennett 2004; Kennett 2011).

<sup>80</sup> As Spence notes, signals are "observable characteristics attached to the individual that are subject to manipulation by him" (Spence 1973, 357).

signaling the blamer's normative commitments. Moreover, blame is costly enough to be hard to fake, and thus it is also a reliable source of information about the blamer's normative commitments and normative competence. In order to send honest signals in the face of a violation of a norm, the blamer must have interiorized the relevant norms. And, finally, blamers stand to gain long-term benefits from their honest signaling, such as the possibility of being recognized as normatively competent members of a given community. All these considerations motivate the choice of blame as a reliable signal of agents' normative competence and normative commitments.

In this section, I have raised three objections to the theory of blame as a form of signaling. By addressing these three concerns, I have argued that the main function of blame is to signal the blamer's normative commitments and normative competence. In the next chapter, I will discuss a possible application of this account of blame. My aim is to show that the functional theory of blame as a form of signaling could help us address concerns about the notion of non-moral blame.

## CHAPTER FIVE

### NON-MORAL BLAME AS A FORM OF SIGNALING

The aim of this chapter is to apply the insights given by the signaling account of blame to explain cases of non-moral blame. As I have already noted in §4, Chapter One, most contemporary accounts of blame focus on *moral* blame; as a result, the vast majority of both content-based theories and functional theories struggle to account for cases of non-moral blame. When they do not struggle to account for cases of non-moral blame, they simply avoid discussing them. This chapter attempts to fill this gap in the contemporary literature on blame. In order to do so, I will refer to the signaling account of blame that I have defended in the previous chapter.

A great advantage of the signaling account of blame over other theories is its capability to explain cases of non-moral blame. To be more precise, what I consider as an important feature of the signaling account of blame is its ability to provide a *unifying explanation* to both moral and non-moral blame. At its core, the main function of blame is signaling the blamer's normative commitments. In turn, a blamer might be committed both to moral and non-moral norms. Not only the signaling account of blame is able to account for the data points of blame I have identified in §5, Chapter One; it is also able to provide a unifying explanation of our blaming interactions. Whether it is moral or non-moral, blame serves the same function – signaling the blamer's normative competence and normative commitments.

Throughout this dissertation I have claimed on multiple occasions that our actual practices include both cases of moral blame and cases of non-moral blame. Even if intuitively plausible, this claim has been recently challenged by B. Matheson and P.E. Milam, who argue against the permissibility of non-moral blame (Matheson and Milam 2021). It should be

noted that these authors *do not claim* that non-moral blame is *conceptually* impossible.<sup>81</sup> Matheson and Milam challenge non-moral blame at the level of our actual practices. As I will explain in this chapter, according to these authors most alleged cases of non-moral blame actually are cases of moral blame. Moreover, the vast majority of instances of non-moral blame are morally impermissible. I find it important to take this challenge seriously, for it questions the plausibility of my claims about our interpersonal and intrapersonal practices of non-moral blame. For this reason, this chapter also provides an answer to the challenge issued by Matheson and Milam.

This chapter is divided into three sections. In §1, I present the challenge against non-moral blame issued by Matheson and Milam. This challenge is constituted by two claims. According to the first claim, when scrutinized with attention, our actual practices include way less cases of non-moral blame than we may initially think. The aim of §1 is showing that this claim does not threaten my characterization of our practices of non-moral blame. In §2, I present and discuss the second part of the challenge to non-moral blame issued by Matheson and Milam. According to the second claim of their challenge, most of the (few) genuine instances of non-moral blame are morally impermissible, e.g., there are overriding moral reasons not to non-morally blame agents. In §3, I show that we have compelling reasons to think that non-moral blame parallels moral blame under the relevant aspects. Thus, a convincing theory of blame should be able to explain both moral and non-moral blame in a unifying way. Finally, I argue that the signaling account of blame is able to provide such a unifying explanation.

---

<sup>81</sup> I have discussed the conceptual challenge against non-moral blame in §4, Chapter One.



## §1 The Case against Non-Moral Blame: The Moralizing Trap

In this section, I present and discuss the first claim that constitutes the challenge against non-moral blame issued by Matheson and Milam (Matheson and Milam 2021).

Non-moral blame seems to be a familiar part of both our interpersonal and intrapersonal experiences. Sport enthusiasts *blame* players and teams when they fail to achieve a victory; citizens *blame* each other for their poor voting choices; an artist may be *blamed* for her bad performance; chess players may *blame themselves* for their blunders; scientists *blame* their colleagues when they fail to exercise their intellectual skills. These reactions seem parallel to cases of moral blame, except for the fact that the person being blamed has not violated a *moral* standard. In all the above cases, the agent being blamed has performed an action that does not meet a normative standard, where this standard is not moral. Scientists and voters who do not exercise their intellectual capabilities can be *epistemically blamed* for their poor reasoning; artists can be *aesthetically blamed* for their performance; athletes can be said to be *skillfully blamed* for their poor display of athletic skills. Thus, at least at an intuitive level, these all are cases of non-moral blame.

Despite the intuitive plausibility of the existence of forms of non-moral blame, our non-moral blaming practices have been recently questioned by Matheson and Milam. In their *Case against Non-moral Blame*, these two authors argue that most instances of alleged non-moral blame are – upon inspection – cases of misidentified *moral* blame (Matheson and Milam 2021, 205). Moreover, most of the (rare) instances of genuine non-moral blame are impermissible from a normative point of view. This challenge against non-moral blame is a serious threat to any account of blame that aims at giving a unifying explanation to both cases of moral blame and cases of non-moral blame. The signaling account of blame provides such a unifying explanation of blame. Thus, I think it is important to address the challenge against non-moral blame issued by Matheson and Milam. In this section, I address the first

part of their challenge, according to which “many (perhaps most) cases that appear to be non-moral blame are actually moral blame” (Matheson and Milam 2021, 205). The authors use the expression “Moralizing Trap” to describe this problem.

To understand how the Moralizing Trap works, consider again one of the cases of alleged non-moral blame that I have mentioned in the opening paragraph of this section. Consider – as Matheson and Milam also do – sport enthusiasts who blame players for their faulty performances. One may think that the reaction of sport fans counts as blame because they are reactions to *violations of non-moral standards*. Players may perform poorly, and as a result they get blamed for not having met athletic standards.<sup>82</sup> This intuitive diagnosis of a case of non-moral blame – which may be extended to other non-moral domains and non-moral standards – is the target of the Moralizing Trap. According to Matheson and Milam, if we put under philosophical scrutiny our intuitions about cases of non-moral blame, we would find that “it is actually [agents’] *moral failure* that we object to and [it is] their culpable *failure to meet a moral obligation* that renders blame permissible in these cases” (Matheson and Milam 2021, 205; emphasis added). To strengthen their position, they point out to cases such as Maradona’s “Hand of God”, where football fans blamed Maradona for having touched the ball with his hand. It can be argued that this is a case of non-moral blame, where blamers blame an agent for having violated a football-related norm. Upon inspection, however, it would be easy to see that this is a case of *moral* blame: football fans blamed Maradona for having *cheated* and having stolen the English chance for a World Cup victory (Matheson and Milam 2021, 205). In my opinion, however, this example is not really effective. I think that even proponents of non-moral blame – among which I consider myself – believe that this particular case should be classified as an instance of moral blame. After all, it seems plausible to consider cheating a case of moral wrongdoing. Thus, it can be plausibly objected to

---

<sup>82</sup> Athletic standards can be seen as a subset of skill-related standards.

Matheson and Milam that they have to provide more compelling examples to convince us that most instances of alleged non-moral blame are cases of moral blame. This does not mean that proponents of non-moral blame can always easily avoid the Moralizing Trap. In the remaining of this section, I present two arguments that threaten our practices of alleged non-moral blame.

In order to claim that most instances of non-moral blame are, upon inspection, cases of moral blame, Matheson and Milam propose two arguments. According to the first argument, non-moral blamers, when confronted about the reasons they had to blame an agent, tend to fall back on *moral* reasons (Matheson and Milam 2021, 205). Thus, according to Matheson and Milam, genuine non-moral blame is a much less widespread phenomenon than we may think. To illustrate this point, consider the following case. You may blame a musician for her poor artistic performance. Absent further details about this scenario, one could think that your blame should count as non-moral blame.<sup>83</sup> However, when confronted about your reasons to blame the musician, you may find yourself thinking that she should have practiced more, that her poor performance reveals her lack of care about her duty, or even that she has violated a professional standard. These may count as moral reasons to blame the musician.<sup>84</sup> If this is correct, then you were actually *morally blaming* the musician. How should we evaluate this argument put forward by Matheson and Milam? I find it plausible that, occasionally, blamers fall back on moral reasons when confronted about the reasons they had to blame an agent. I do not think that it is particularly problematic to accept that sometimes blamers misidentify moral blame for non-moral blame. However, I do not think that Matheson and Milam have the resources to argue in favor of a *stronger* claim,

---

<sup>83</sup> Plausibly, this would count as a case of aesthetic blame. However, in this chapter I tend not to qualify the normative domain of our practices of non-moral blame. After all, Matheson and Milam challenge the core notion of non-moral blame; they do not argue against the plausibility of, say, epistemic or aesthetic blame (Matheson and Milam 2021). Thus, I prefer not to focus on a specific kind of non-moral blame.

<sup>84</sup> Note that critics of non-moral blame would still need to prove that violations of professional standards count as moral wrongdoings.

according to which *most* blamers would – upon reflection – fall back on moral reasons. Thus, I believe that proponents of non-moral blame should simply bite the bullet and admit that sometimes we are wrong about our intuitions concerning cases of (alleged) non-moral blame. At the same time, proponents of non-moral blame should argue that the burden of proving that *most* blamers would fall back on moral reasons is on Matheson and Milam.<sup>85</sup> It may well be true that – upon inspection – some cases of alleged non-moral blame are cases of misidentified moral blame; however, not every blamer would fall back on moral reasons for blaming an agent. So, genuine non-moral blame seems safe against the Moralizing Trap.

I find the second argument provided by Matheson and Milam in favor of the Moralizing Trap more compelling than the one that I have discussed above. According to the second argument, it is the *moral* features of a particular case that provide the blamers with reasons to blame someone. Thus, even if our intuitions suggest identifying the blaming interaction as a case of non-moral blame, we should classify such interaction as moral blame (Matheson and Milam 2021, 205-207). Another way to illustrate the same point is thinking about how blamers' attitude would change – *if* they would change – after having stripped away the moral features of a particular case. In order to understand Matheson's and Milam's argument against non-moral blame, consider these two similar scenarios.

*Bomb Disposal Expert:* A bomb disposal expert is trying to defuse a bomb. If she fails to defuse the bomb, the safety of hundreds of people would be at risk. The bomb disposal expert fails to defuse the bomb, and as a result some civilians get hurt.

*Bomb Disposal Novice in Training:* A bomb disposal novice is training in a military camp. She is practicing how to defuse a bomb. If she fails to defuse the bomb, the consequences would

---

<sup>85</sup> Matheson and Milam argue that “resorting to moral grounds for blame seems common”, but they do not provide any empirical data to support this claim (Matheson and Milam 2021, 205).

not be harmful to anyone. The bomb disposal novice fails to defuse the bomb, without harmful consequences.

Now, if we were prone to blame the bomb disposal expert from the first case, but not the bomb disposal novice from the second case, our reasons for blaming would be moral reasons. The second case, *Bomb Disposal Novice in Training*, does not present any morally relevant feature. It seems, then, that absent morally relevant features we are not prone to blame agents. If this is so, critics of non-moral forms of blame seem right in claiming that there is something wrong in thinking that our actual blaming practices include non-moral blame. The analysis I have just proposed of the comparison between *Bomb Disposal Expert* and *Bomb Disposal Novice in Training*, however, is misleading. The defender of non-moral blame should note that the object of blame in *Bomb Disposal Expert* is *the consequence* of failing to defuse the bomb. Instead of blaming the bomb disposal expert for having *failed to defuse the bomb*, moral blame has as its object the endangerment of other people's lives. Endangering other people's lives is absent in the second case, and this should explain the absence of a blaming reaction. But is this all we can say about the comparison between the two cases? I do not think so. In both cases, we can blame the bomb disposal expert or novice for her *having failed to defuse the bomb*, which is not – *per se* – a moral failure. Defusing a bomb, just as solving a mathematical problem, is a skill. And it seems that we can react by blaming agents who fail to meet justifiable expectations related to their skills (Björnsson 2022). Think of the reaction that the bomb disposal novice trainer would have upon finding out that she has failed to perform her task. The trainer would blame the novice for not having met a standard related to a skill. In other words, the trainer would blame the novice for not having met a non-moral standard. My contention is that in *Bomb Disposal Expert* we morally blame the bomb disposal expert for endangering other people's lives, but we can also non-morally blame her for her skill-related failure in missing the opportunity to defuse the bomb. In *Bomb*

*Bomb Disposal Novice in Training*, being moral features absent from the description of the case, we can non-morally blame the bomb disposal novice for her skill-related shortcoming, which is a non-moral failure. So, *pace* Matheson and Milam, blame is not absent in the second case: it is moral blame which is absent. But this is coherent with the description of the case. As we know, in the description of *Bomb Disposal Novice in Training* moral features are left out on purpose.

To sum up, we can see that both arguments in favor of the Moralizing Trap fail to convince us that most instances of alleged non-moral blame are, upon inspection, cases of moral blame. If so, I think we are motivated to find a theory of blame that is able to account for cases of non-moral blame as well as for cases of moral blame. The signaling theory of blame, as I will argue in §3, can meet this *desideratum*. However, before arguing that the signaling account of blame provides a compelling explanation of cases of non-moral blame, I want to address another challenge to proponents of non-moral blame. After all, showing that our actual blaming practices include cases of non-moral blame may be insufficient. What if, as Matheson and Milam claim, genuine non-moral blame is all-things-considered impermissible? If this is so, then maybe we should always avoid blaming ourselves and others for their non-moral failures. In the next section, I will explore this possibility.

## §2 Is Non-Moral Blame Impermissible?

In the previous section, I have shown that – despite the threat posed by the Moralizing Trap – our blaming practices often include cases of genuine non-moral blame, e.g., blame that has as its target an agent who has not violated a moral norm. This may not be sufficient to fully vindicate our practice of non-morally blaming each other. In their *Case against Non-Moral Blame*, Matheson and Milam raise a second challenge directed to proponents of non-moral blame. According to Matheson and Milam, even if our actual practices include cases of

genuine non-moral blame, it could be argued that non-moral blame is hardly ever (perhaps, never) all-things-considered morally justified (Matheson and Milam 2021, 207-209). This means that there are overriding moral reasons to avoid non-morally blaming other agents.<sup>86</sup> In this section, I argue that Matheson and Milam are wrong in thinking that non-moral blame is impermissible. Let me introduce my argument with an example.

Consider the case of Jim. Jim is not a professional guitar player, but he loves playing with his friends during the holidays. Summer comes, and he wants to learn how to play a new song on his guitar. The song he is trying to learn is not too difficult to play, but neither is it trivial. As it is, the new song is entirely within Jim's abilities. He practices for some time, until he thinks he is ready. While he practices at home, he manages to play even the most difficult sections of the new song. After some weeks Jim is ready to play the new song with his group of friends. During a weekend in July, he meets with his friends, and they start playing. When they play the new song, however, Jim misses some notes during his solo. This failure is not due to some unfortunate circumstances, nor has Jim failed to prepare adequately to play the new song with his group of friends. Jim's friends do not take this failure to be of particular relevance: they just carry on playing the song with a smile on their face. They are not laughing at Jim, nor they think that he is a mediocre guitar player: they just take his mistake lightheartedly. Jim, however, blames himself for his own mistake. He knows that he has not met a standard he holds himself up to, and his reaction might be plausibly described as non-moral blame directed to himself. Now, imagine that in order to improve his guitar skills and not fail again, Jim decides to hire a private guitar teacher, Tom. Tom is a professional guitar player and a guitar teacher, and Jim pays Tom for his services. As in the previous scenario, Jim wants to learn a new song on his guitar. With the help of Tom, Jim improves a lot and feels ready to play the new song. He has not neglected his practice sessions

---

<sup>86</sup> Note that this argument refers to what is often labeled as the Ethics of Blame. See §2, Chapter One for an overview of the differences between the nature of blame and the Ethics of Blame.

and he has not been lazy. When he feels ready, he plays the new song together with his private teacher. However, he misses again some notes during the solo. When the song ends, Jim blames himself like in the original scenario. This time, however, he also gets blamed by his private guitar teacher. Are there any reasons to maintain that it is all-things-considered impermissible, for Tom, to blame Jim? I think that critics of non-moral blame would argue that such reasons can be found. However, I maintain that such reasons are not strong enough, and they should be rejected. Thus, I take Tom's non-moral blame to be all-things-considered justifiable.

In this section, I consider two arguments against all-things-considered permissible non-moral blame. Both arguments conclude that, in cases of non-moral failures, forms of *criticism* or *negative evaluation* are to be preferred instead of blame. Criticism or negative evaluation are to be preferred over blame for moral reasons – that is, there are moral reasons to argue that non-moral blame is impermissible. One way of arguing against the permissibility of non-moral blame is claiming that non-moral blame is always disproportionate. Blame is disproportionate when it exceeds what is called for. Being harshly blamed for minor mistakes is often considered impermissible. This is true also in the moral domain. A friend of mine may promise to show up for a coffee tomorrow. If he is, say, ten minutes late, it would be impermissible for me to harshly blame him for his (minor) lack of regard toward me. For example, it would be disproportionate to shout at him and completely avoid any kind of interaction with him for the next month.<sup>87</sup> This is because, in treating my friend in such a way, I would harm him in a way that exceeds the negative moral value of his action. A critic of non-moral blame might think that blaming someone for her non-moral failure is always disproportionate, because blame typically involves some kind of harsh treatment of the blamee. This, however, is false. First, it must be noted that not every instance of blame is an

---

<sup>87</sup> This depends on the description of the situation. Generally, being ten minutes late does not justify harsh treatment. However, if stakes are high enough, harsh blame might be a proportional reaction.



instance of overt or expressed blame. Thus, blame may not involve any kind of *treatment* at all. Second, and more important, blame does not have to be necessarily harsh. Matheson and Milam write that “it seems unreasonably harsh for a critic to describe a piece of art as ‘tasteless hack work’ [...], or for a Reddit contributor to insult and swear at Google’s software engineers in response to a Gmail update” (Matheson and Milam 2021, 210). As far as I am concerned, I think they are correct: in cases such as those they describe, insulting artists and software engineers is disproportionate and thus impermissible. This, however, does not mean that we shall conclude that non-moral blame is all-things-considered unjustifiable *tout court*. It seems to me that at least certain non-moral blaming responses can be proportionate to non-moral failures. I cannot see why a chess player should avoid blaming herself for her mistakes on the chessboard by feeling a (perhaps mild) urge to kick herself for that mistake. And, as far as other-directed blame is concerned, I fail to see why Jim’s private guitar teacher should avoid blaming Jim because of blame being disproportionate. Clearly, Jim’s private teacher cannot permissibly insult Jim and demean him for getting some notes wrong during a solo. This reaction would be impermissible. The guitar teacher, however, does not need to blame Jim by insulting him. He could feel a mildly negatively valued reactive attitude towards Tom for failing to put into practice his teachings and not meeting his justifiable expectations. More likely, Tom would not even feel a blaming emotion, but he can react by mildly reproaching Jim. But this will not be disproportionate. Quite the contrary: it is plausible to think that Jim expects Tom to react in such a way. What would be wrong, here, is insulting Jim for his mistake. I maintain that there is nothing morally wrong, for Jim’s guitar teacher, to blame Jim for his mistake. What is at stake, here, is not blame *per se*, but the involvement of forms of harsh treatment against the person being blamed. However, I fail to see a necessary link between blame and harsh treatment. I suspect that critics of non-moral blame who think that non-moral blame is all-things-considered impermissible because it is disproportionate fail to see the distinction between harsh treatment and blame. As I have

argued, blame – whether it is moral or non-moral – does not necessarily involve insults or harsh treatment. In addition, it can be argued that certain reactions to non-moral shortcomings play a role in influencing (e.g., improving) the conduct of the person being blamed *precisely because* they are felt as forms of *blame*, and not as forms of criticism.

Another argument against the all-things-considered permissibility of non-moral blame involves considerations about the outcomes of our blaming activities. According to such an argument, blame is all-things-considered impermissible if it produces bad enough consequences (Matheson and Milam 2021, 210-211). If blaming an agent has overall terrible effects, then it is (morally) wrong to blame that agent. Even if we leave unspecified what the bad outcomes could be, I agree with Matheson and Milam when they claim that the outcomes of blame are not only consequentialist concerns (Matheson and Milam 2021, 211). However, I do not agree with them when they claim that “the badness of blame is especially important in the case of non-moral blame because the need to engage in non-moral blame seems less pressing than moral blame” (Matheson and Milam 2021, 210). It is not entirely clear to me what the “need to engage in non-moral blame” is. Moreover, I cannot see why this need should be less pressing in cases of non-moral blame. I suspect that what is at work here is a moralized version of non-moral blame. Matheson and Milam may think that it is sometimes required to blame someone for his or her moral failures. This could explain the “need” to engage in blaming activities.<sup>88</sup> For example, upon learning that our neighbor is a racist, we may feel it is morally required for us to blame him and avoid further interactions with him. Non-moral blame, on the contrary, does not seem to be *morally required*. However, think again of the relationship between Jim, the guitar amateur, and Tom, his private teacher. Because of the nature of this relationship, we may think that Tom is required to blame Jim for his

---

<sup>88</sup> The need of blaming someone may also be explained as an uncontrollable urge. I am not satisfied by this explanation. First, even emotional accounts of blame often maintain that reactive attitudes are not knee-jerking responses, lacking any conceptual component. Second, if the needs of blaming were truly uncontrollable, then it would be impossible for the prospective blamer to refrain from blaming and opting for criticism.

mistake. We may think so if we consider the positive outcomes that this blaming activity would have – after all, Jim hired a private teacher precisely because he wants to improve. After being blamed by his teacher, Jim focuses even more on practicing his musical skills. Critics of non-moral blame, however, may argue that the positive outcomes of such blame would be outweighed by the even more positive outcomes of criticism. I think that this largely depends on how we describe scenarios involving the all-things-considered permissibility of non-moral blame. I have two considerations on this. First, note that criticism may bring about negative outcomes. I fail to see why, *per se*, criticism would be less harmful than blame. It seems clear enough that much depends *on the way* in which we criticize agents for their failures. Criticism is not necessarily less harmful than blame. On the contrary, it is potentially more harmful: as said, much depends on the details. Second, it can be argued that blame is a reaction to *agential* shortcomings. Criticism, on the other hand, may be used to point out negative features of an *action*.<sup>89</sup> Typically, these features also explain why other courses of action would have been preferable. Thus, criticism does not necessarily involve an agential shortcoming. Depending on the situation, the desired outcomes may be achieved either by non-moral blame or by criticism. This suggests that it is a mistake to believe that, in order to avoid bad outcomes, criticism should always replace blame in cases of non-moral failures.

To sum up, it seems that there are no sufficiently compelling reasons to conclude that most of our non-moral blame should be deemed impermissible. As a result, I do not think that we should revise our practices in order to expunge non-moral blame from them. In the introductory section of this chapter, I claimed that our blaming practices include cases of non-moral blame. Then, I presented a challenge against non-moral blame. According to

---

<sup>89</sup> G. Björnsson aptly distinguishes between normatively negative outcomes due to agential shortcomings and outcomes that “just happen” to be normatively negative (Björnsson 2022). Blame would be out of place in case the outcome of an action performed by an agent is not due to an agential shortcoming on the part of the agent.

this challenge, most (perhaps every) instance of alleged blame are instances of moral blame. Moreover, from a normative point of view, cases of genuine non-moral blame are all-things-considered impermissible. In both §1 and §2, I have argued that the challenge against non-moral blame can be resisted. In the following section, I show that the signaling account of blame can provide a compelling explanation of non-moral blame.

### **§3 Non-moral Blame and the Signaling Theory of Blame**

In this section, I will argue that the signaling account of blame that I have defended in the previous chapter provides a straightforward explanation of cases of non-moral blame. My argument is constituted by two steps. First, I argue that our practice of non-moral blame parallels our practice of moral blame in relevant respects. Then, I claim that the signaling account of blame can be employed to explain cases of moral blame just as it can be employed to explain cases of non-moral blame. In both cases of moral and non-moral blame, the function of blame is signaling the blamer's normative commitments.

Let me start by highlighting the parallel between moral blame and non-moral blame. To strengthen my argument against the challenge to non-moral blame issued by Matheson and Milam, I follow their own characterization of the similarities between moral and non-moral blame (Matheson and Milam 2021, 204-206). First, they note that both moral and non-moral blame involve more than the mere recognition of a causal link between an agent and a normatively substandard action performed by such agent. This suggests that blame includes an evaluative dimension, recognized by most authors (Coates and Tognazzini 2013). Moreover, as normatively competent beings, we can distinguish between judgments of blameworthiness and active blame. As normatively competent beings, we are able to understand that one thing is to believe that blame is a fitting or appropriate response to the

behavior of an agent,<sup>90</sup> while another thing is to actively blame such agent. Additionally, blamers hold the target(s) of their blaming activities to certain normative standards. When we blame an agent, we react to an agent's failure to meet certain normative standards. As Matheson and Milam propose, "If moral blame is a response to culpable moral failures (e.g. breaking a promise), then non-moral blame is a response to a culpable failure to meet a non-moral standard (e.g. failing to make a penalty kick)" (Matheson and Milam 2021, 202). Thus, we can appreciate that even critics of non-moral blame admit that our practice of non-moral blame parallels our practice of moral blame in many respects.

In addition, recall the list of data points of blame that I have presented in §5, Chapter One. Are there any reasons for restricting the data points of blame to moral blame? I do not think so. After all, both moral and non-moral blame can be overtly expressed or kept private; both moral and non-moral blame can directly address the person being blamed or being expressed towards third parties (as in cases of third-party blame); blamers may blame themselves both morally and non-morally; both moral and non-moral blame can be "dispassionate" (e.g., not accompanied by a blaming emotion). Finally, both moral and non-moral blame can target distant or dead persons. These are all additional reasons to accept the parallel between moral blame and non-moral blame.

Despite the similarities between moral and non-moral blame, the plausibility of non-moral blame can still be challenged from a theoretical point of view. For example, it could be objected that in order to be *blameworthy*, agents must be responsible. And, the argument goes on, agents cannot be non-morally responsible. If this is true, it might be concluded that the parallel between moral blame and non-moral blame breaks down, because if agents cannot be non-morally responsible, they cannot even be non-morally blameworthy.

---

<sup>90</sup> Some prefer a stronger interpretation of judgments of blameworthiness and refer to the notion of desert. According to such interpretation, blameworthy agents *deserve* blame. However, I prefer referring to fittingness or appropriateness, where "appropriate" does not mean "*morally* appropriate".

However, I think we have strong reasons to reject this argument. First, note that this is a conceptual point. Even if it were true that non-moral blameworthiness is impossible, one would still have to explain why our actual interpersonal practices seem to include cases of non-moral blame, e.g., cases in which someone blames an agent for his or her culpable non-moral failure.<sup>91</sup> Second, it should be noted that most theories of normative responsibility identify conditions of responsibility that can be met by normatively competent agents regardless of whether imputable actions possess moral significance (Wolf 2015). Normative responsibility, in this sense, includes both moral and non-moral responsibility (Peels 2017).<sup>92</sup> To conclude, I think that pointing at alleged asymmetries between moral and non-moral responsibility does not help in arguing against non-moral blameworthiness and non-moral blame.

Critics of non-moral blame, however, can insist that there is a crucial difference between moral and non-moral blame. As I have already mentioned, Matheson and Milam claim that moral blame is a response to culpable moral failure (Matheson and Milam 2021, 202). Crucially, they also believe that blame is a reaction to the violation of an *obligation*. If blame is a reaction to the violation of an obligation, however, proponents of non-moral blame face a challenge. While most of us would acknowledge the existence of moral obligations, the existence of non-moral obligations is debatable. If blame requires that an agent has violated an obligation, and if there are not non-moral obligations, it follows that all blame is *moral* blame.<sup>93</sup> In their case against non-moral blame, Matheson and Milam refer to this argument in order to rule out the possibility of genuine non-moral blame (Matheson and Milam 2021, 208). However, I think that this argument rests on wrong premises. First,

---

<sup>91</sup> A possible solution would include the need to revise our interpersonal practices in order to systematically avoid cases of non-moral blame; however, I find this solution unappealing.

<sup>92</sup> See also Nelkin 2020 for an attempt to build a unifying framework for both moral and non-moral responsibility (Nelkin 2020).

<sup>93</sup> Another way to see this is arguing that moral wrongdoing is a necessary condition of blame.

contrary to Matheson and Milam, I think that the second premise of the above argument, which states that there are not non-moral obligations, is highly controversial – if not entirely mistaken. Many epistemologists, for example, would argue that there are epistemic obligations concerning our beliefs (Feldman 1988). The aesthetic domain, moreover, seems to admit the presence of aesthetic obligations (Archer and Ware 2018). For the sake of the argument, however, we can accept this premise. I argue that, even if we accept that there are not non-moral obligations, we should reject that blame requires the violation of an obligation. Let me explain why.

Most philosophers argue that blame has a distinctive force or sting (Hieronymi 2004). As Matheson and Milam explicitly claim, blame differs from mere negative evaluation because “the sting of blame is part of its purpose. This is so even if one doesn’t intend for blame to hurt or if the blame is private” (Matheson and Milam 2021, 201). I suspect that Matheson and Milam believe that violating an obligation is necessary for blame because they think that violations of an obligation are the only things that warrant the characteristic force of blame. It seems to me that critics of non-moral blame think that *only moral wrongdoing* lends blame its distinctive force or sting.<sup>94</sup> However, this is a mistake. Thanks to the theory of blame I have defended in the previous chapter, we can account for the distinctive force of blame without having to refer to violations of an obligation.

Let me briefly recall the distinctive features of the signaling account of blame. According to the signaling account of blame, the function of blame is signaling the blamer’s normative competence and normative commitments. More precisely, blame acts as a *costly signal* of the blamer’s normative commitments. My contention is that the distinctive force or sting of blame is sufficiently *explained by the costs* associated both with blame and with being a

---

<sup>94</sup> Incidentally, this force or sting is also what it is thought to let us distinguish blame from criticism (Bennett 2013; Scanlon 1998, Ch. 6; Smith 2008).

blaming agent. In order to be a reliable signal of one's normative commitments, blame has to be costly; otherwise, free riders would exploit the practice of blaming in order to fake their normative commitments and to get the benefits associated with being part of the blaming community. It is the cost associated with blame that lends blame its distinctive force or sting, and not the violation of an obligation. As I have argued in §3, Chapter Four, blamers need to undergo a process of interiorization of the norms they are committed to. This process is particularly costly, and it triggers blaming responses as a reaction to violations of the internalized norms. This is what lends blame its characteristic sting. Thus, I believe that the signaling account of blame possesses the resources to argue against the claim according to which blame requires the violation of an obligation.

In addition, both our interpersonal and intrapersonal blaming practices lend support to the thesis according to which blame – being it moral or non-moral – is, at its core, a form of signaling our normative competence and commitments. Think of the reaction of a tennis professional player upon realizing that her shot failed to reach the putting green. She may feel the urge to kick herself for her mistake, and she can even throw her golf as a sign of her frustration (Shoemaker and Vargas 2021, 593). All the elements which typically characterize blame are present: an agent culpably failed to meet a normative standard; someone reacts to this performance by addressing the agent; the reaction is emotionally toned (even if it could have been dispassionate); the reaction goes beyond a form of mere evaluation, even if it has an evaluative component). My contention is that what unifies this practice with cases of moral blame is that blame discharges the same function – that is, signaling the blamer's normative commitments.

In this section, I have highlighted the many similarities between moral and non-moral blame. My conclusion is that we have no particular reasons to build two different accounts of blame in order to explain moral blame and non-moral blame as separate phenomena. The



signaling account of blame provides a unifying explanation of blame and of our blaming practices. Independently from the fact that a particular instance of blame counts as moral or non-moral, the function of blame is the same – blame signals the blamer’s normative competence and normative commitments.

## CONCLUSION

Two were the main aims of this dissertation. First, to offer a taxonomy of the contemporary theories of the nature of blame. These theories attempt to answer the following questions: what is blame, and which human interactions count as blaming interactions? The second and the third chapter of this dissertation achieved this aim by offering a reconstruction of the contemporary debate surrounding the nature of blame. While presenting both content-based and functional accounts of blame, I have also argued that most contemporary theories of blame fail at giving a unifying explanation to our blaming practices. This consideration, however, should not discourage us from continuing to pursue an answer to the main questions about the nature of blame.

In the fourth chapter of this dissertation, I have defended a functional account of blame that offers a unifying explanation of our blaming practices. According to such an account, the function of blame is signaling the blamer's normative competence and normative commitments. Not only the signaling account of blame does better than its competitors at accounting for the various data points of blame, but it is also able to give a unifying explanation to both moral and non-moral blame. I have applied the signaling model of blame to non-moral blame in the fifth and final chapter of this dissertation. I consider this to be an important advantage of the signaling account of blame over most of the other theories of blame.

Theories of the nature of blame tend to stay neutral about normative questions that surround blame. This does not mean that the normative structure of our blaming activities should be ignored. The so-called Ethics of Blame constitutes an important part of any theory of blame. In this dissertation, I did not discuss the Ethics of Blame; instead, I proposed a possible solution to the problem of the nature of blame. The signaling account of blame that I have defended in this dissertation is well-equipped to address normative questions about

our blaming practices. I take this to be a promising starting point for future research on blame.

## REFERENCES

- Archer, A. and Waren, L. 2018, “Beyond the Call of Beauty: Everyday Aesthetic Demands Under Patriarchy”, *The Monist*, 101.
- Archetti, M. 2000, “The Origin of Autumn Colours by Coevolution”, *Journal of Theoretical Biology*, 205 (4): 625-630.
- Arpaly, N. 2006, *Merit, Meaning, and Human Bondage: An Essay on Free Will*, Princeton University Press.
- Beardsley, E.L. 1969, “A Plea for Desert”, *American Philosophical Quarterly*, 6(1): 33-42.
- Beardsley, E.L. 1979, “Blaming”, *Philosophia*, 8(4): 573-583.
- Bell, M. 2013, “The Standing to Blame: A Critique”, in Coates & Tognazzini 2013: 263-281.
- Bennett, C. 2013, “The Expressive Function of Blame”, in Coates & Tognazzini 2013: 66–83.
- Bird, R. B., Smith, E. and Bird, D. W. 2001, “The hunting handicap: Costly signaling in human foraging strategies”, *Behavioral Ecology and Sociobiology*, 50, 9–19.
- Bonicalzi, S. and De Caro, M. 2013, *Free Will: Thirty Points of View*. Methode: Analytic Perspectives, Università di Torino, 2(3).
- Björnsson, G. 2017, “Explaining Away Epistemic Skepticism about Culpability”, in Shoemaker 2017a: 141–162.
- Björnsson, G. 2022, “Blame, deserved guilt, and harms to standing”, in Carlsson 2022.
- Chislenko, E. 2021, “Causal Blame”, *American Philosophical Quarterly*, 58(4): 347-358.
- Clarke, Randolph, 2009, “Dispositions, Abilities to Act, and Free Will: The New Dispositionalism”, *Mind*, 118(470): 323–351.

- Coates, D.J. and Tognazzini N. 2012, “The Nature and Ethics of Blame”, *Philosophy Compass*, 7/3: 197-207.
- Coates, D.J. and Tognazzini N. 2013a, “The Contours of Blame”, in Coates and Tognazzini 2013b: 3–26.
- Coates, D.J. and Tognazzini N. (eds.), 2013b, *Blame: Its Nature and Norms*, New York: Oxford University Press.
- Coates, D.J. and Tognazzini N. 2021, “Blame”, "Blame", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/blame/>.
- Coates, D.J. 2016, “The Epistemic Norm of Blame,” *Ethical Theory and Moral Practice*, 19: 457–473.
- Darwall, S. 2006, *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press.
- Driver, J. (2016), “Private Blame”. *Criminal Law and Philosophy*, 10 (2): 215-220.
- Dawkins, R. & Krebs, J.R. 1978, “Animal signals: information or manipulation”, in *Behavioural ecology: an evolutionary approach*, ed. J. R. Krebs & N. B. Davies, 282-309. Oxford: Blackwell Scientific.
- Eaton, M. 2008, “Aesthetic Obligations”, *Journal of Aesthetics and Art Criticism*, 66:1.
- Enoch, D. 2012, “Being Responsible, Taking Responsibility, and Penumbral Agency”, in *Luck, Value, and Commitment: Themes from the Ethics of Bernard Williams*, Ulrike Heuer and Gerald Lang (eds.), Oxford: Oxford University Press, 95–132.
- Eriksson, L. 2019, “Social Norms as Signals”, *Social Theory and Practice*, 45(4): 579-599.

- Eshleman, A. 2014, “Worthy of Praise: Responsibility and Better-than-Minimally-Decent Agency”, in Shoemaker and Tognazzini 2014: 216–242.
- FitzPatrick, W.J. 2008, “Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge”, *Ethics*, 118(4): 589–613.
- Fonnesu, L. 2013, “Genealogie della responsabilità”, in *Quanto siamo responsabili? Filosofia, neuroscienze e società*, Torino: Edizione Codice, 6-23.
- Franklin, C.E. 2011, “Valuing Blame”, in Coates & Tognazzini 2013: 207-223.
- Fricker, M. 2016, “What’s the Point of Blame? A Paradigm Based Explanation”, *Noûs*, 50(1): 165–183.
- Friedman, M., 2013, “How to Blame People Responsibly”, *The Journal of Value Inquiry*, 47: 271–284.
- Glannon, W. 1997, “Psychopathy and Responsibility”, *Journal of Applied Philosophy*, 14(3): 263–275.
- Graham, P. 2014, “A Sketch of a Theory of Moral Blameworthiness”, *Philosophy and Phenomenological Research*, 2, 388-409.
- Greenspan, P.S. 2003, “Responsible Psychopaths”, *Philosophical Psychology*, 16(3): 417–429.
- Guerrero, A. 2007, “Don’t Know, Don’t Kill: Moral Ignorance, Culpability, and Caution”, *Philosophical Studies*, 136(1): 59–97.
- Hamilton, W. D. and Brown, SP 2001, "Autumn tree colours as a handicap signal", *Proceedings of the Royal Society*, 268 (1475): 1489–1493.
- Haji, I. 1998, *Moral Appraisability*, New York: Oxford University Press.

- Hartman, R.J. 2017, *In Defense of Moral Luck: Why Luck Often Affects Praiseworthiness and Blameworthiness*, New York: Routledge.
- Hieronymi, P. 2001, “Articulating an Uncompromising Forgiveness”, *Philosophy and Phenomenological Research*, 62(3): 529–555.
- Hieronymi, P. 2004, “The Force and Fairness of Blame”, *Philosophical Perspectives*, 18(1): 115–148.
- Hieronymi, P. 2008, “Responsibility for Believing”, *Synthese*, 161(3): 357–373.
- Hieronymi, P. 2014, “Reflection and Responsibility: Reflection and Responsibility”, *Philosophy & Public Affairs*, 42(1): 3–41.
- Irons, W. (1975), *The Yomut Turkmen: A Study of Social Organization among a Central Asian Turkic-Speaking Population*, Ann Harbor: University of Michigan Press.
- Irons, W. (2001), “Religion as a hard-to-fake sign of commitment”. In *Evolution and the capacity for commitment*. R. M. Nesse (Ed.), Russell Sage Foundation, 290–309.
- Kelly, Erin I., 2013, “What Is an Excuse?”, in Coates and Tognazzini 2013b: 244–262.
- Kenner, L. 1967, “On Blaming”, *Mind*, 76(302): 238-249.
- Kennett, J. 2019, “Competence, Attributability, and Blame: Resolving the Responsibility of the Psychopath”, in *Oxford Studies in Agency and Responsibility*, Volume 5, D. Justin Coates and Neal A. Tognazzini (eds.), Oxford: Oxford University Press, 142–164.
- Levy, N. 2003, “Cultural Membership and Moral Responsibility”, *The Monist*, 86(2): 145–163.
- Levy, N. 2005, “The Good, the Bad, and the Blameworthy”, *Journal of Ethics and Social Philosophy*, 1(2): 1–16.

- Levy, N. 2007, “The Responsibility of the Psychopath Revisited”, *Philosophy, Psychiatry, & Psychology*, 14(2): 129–138.
- Levy, N. 2012, “Skepticism and Sanction: The Benefits of Rejecting Moral Responsibility”, *Law and Philosophy*, 31(5): 477–493.
- Luthra, Y. 2016, “Non-rational aspects of skilled agency”, *Philosophical Studies* 173.8: 2267-2289
- Macnamara, C. 2015a, “Blame, Communication, and Morally Responsible Agency”, in *The Nature of Moral Responsibility: New Essays*. Randolph Clarke, Michael McKenna, and Angela M. Smith (eds.), New York: Oxford University Press, 211–236.
- Macnamara, C. 2015b, “Reactive Attitudes as Communicative Entities”, *Philosophy and Phenomenological Research*, 90: 546–569.
- Maibom, H.L. 2008, “The Mad, the Bad, and the Psychopath”, *Neuroethics*, 1(3): 167–184.
- Mason, E. 2015, “Moral Ignorance and Blameworthiness”, *Philosophical Studies*, 172(11): 3037–3057.
- Mason, E. 2019, *Ways to Be Blameworthy: Rightness, Wrongness, and Responsibility*, Oxford: Oxford University Press.
- Mauss, M. 1924, *The gift: Forms and functions of exchange in archaic societies*. London: Cohen and West.
- McGeer, V. 2015, “Building a Better Theory of Responsibility”, *Philosophical Studies*, 172(10): 2635–2649.
- McKenna, M. 2004, “Responsibility and Globally Manipulated Agents”, *Philosophical Topics*, 32(1/2): 169–192. doi:10.5840/philtopics2004321/222



- McKenna, M. 2008, "A Hard-Line Reply to Pereboom's Four-Case Manipulation Argument", *Philosophy and Phenomenological Research*, 77(1): 142–159.
- McKenna, M. 2012, *Conversation and Responsibility*, New York: Oxford University Press.
- McKenna, M. 2013, "Reasons-Responsiveness, Agents, and Mechanisms", in Shoemaker 2013: 151–183.
- Matheson, B. and Milam, P.E. 2021, "The case against Non-moral Blame", in *Oxford Studies in Normative Ethics* 11.
- Nelkin, D. 2008, "Responsibility and Rational Abilities: Defending an Asymmetrical View", *Pacific Philosophical Quarterly*, 89(4): 497–515.
- Nelkin, D. 2011, *Making Sense of Free Will and Responsibility*, New York: Oxford University Press.
- Nelkin, D. 2015, "Psychopaths, Incurable Racists, and the Faces of Responsibility", *Ethics*, 125(2): 357–390.
- Owen, D. 2012, *Shaping the normative landscape*. Oxford: Oxford University Press.
- Peels, R. 2016, *Responsible Belief: A Theory in Ethics and Epistemology*, Oxford: Oxford University Press.
- Polger, T.W. 2019, "Functionalism", *The Internet Encyclopedia of Philosophy*.
- Robichaud, P. and Wieland J. W. (eds.) 2017, *Responsibility: The Epistemic Condition*, Oxford: Oxford University Press.
- Rorty, A. 2010, "Questioning Moral Theories", *Philosophy*, 85(1): 29-46.
- Rosen, G. 2003, "Culpability and Ignorance", *Proceedings of the Aristotelian Society*, 103(1): 61–84.

- Rosen, G. 2004, "Skepticism about Moral Responsibility", *Philosophical Perspectives*, 18: 295–313.
- Sartorio, C. 2007, "Causation and Responsibility", *Philosophy Compass* 2 (5):749–765.
- Scanlon, T. M., 1998, *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
- Scanlon, T.M. 2008, *Moral Dimensions: Permissibility, Meaning, and Blame*, Cambridge, MA: Harvard University Press.
- Sher, G. 2006a, *In Praise of Blame*, New York: Oxford University Press.
- Sher, G. 2006b, "Out of Control", *Ethics*, 116(2): 285–301.
- Sher, G. 2009, *Who Knew? Responsibility Without Awareness*, New York: Oxford University Press.
- Shoemaker, D. 2011, "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility", *Ethics*, 121(3): 602–632.
- Shoemaker, D. 2012, "Responsibility Without Identity", *The Harvard Review of Philosophy*, 18: 109–132.
- Shoemaker, D. (ed.) 2013, *Oxford Studies in Agency and Responsibility*, Volume 1, New York: Oxford University Press.
- Shoemaker, D. 2014, "Psychopathy, Responsibility, and the Moral/Conventional Distinction", in Schramme 2014: 247–74.
- Shoemaker, D. 2015a, *Responsibility from the Margins*, New York: Oxford University Press.
- Shoemaker, D. (ed.) 2015b, *Oxford Studies in Agency and Responsibility*, Volume 3, New York: Oxford University Press.

Shoemaker, D. (ed.), 2017a, *Oxford Studies in Agency and Responsibility*, Volume 4, New York: Oxford University Press.

Shoemaker, D. 2017b, “Response-Dependent Responsibility; or, A Funny Thing Happened on the Way to Blame”, *Philosophical Review*, 126(4): 481–527.

Shoemaker, D. and Tognazzini, N. (eds.), 2014, *Oxford Studies in Agency and Responsibility*, Volume 2: “Freedom and Resentment” at 50, New York: Oxford University Press.

Shoemaker, D. and Vargas, M. 2021, “Moral Torch Fishing: A Signaling Theory of Blame”, *Noûs*, 55: 581-602.

Sliwa, P. 2021, “Reverse-Engineering Blame”, *Philosophical Perspectives*, 33: 200-219.

Smart, J. J. C. 1961, “Free-Will, Praise and Blame”, *Mind*, 70(279): 291–306.

Smart, J.J.C. 1973, “An Outline of a Utilitarian System of Ethics”, in *Utilitarianism: For and Against*, J. J. C. Smart and Bernard Williams, Cambridge: Cambridge University Press, pp. 3–74.

Smith, A. 2005, “Responsibility for Attitudes: Activity and Passivity in Mental Life”, *Ethics*, 115(2): 236–271.

Smith, A. 2008, “Control, Responsibility, and Moral Assessment”, *Philosophical Studies*, 138(3): 367–392.

Smith, A. 2013, “Moral Blame and Moral Protest”, in Coates and Tognazzini 2013: 27–48.

Smith, A. 2015, “Responsibility as Answerability”, *Inquiry*, 58(2): 99–126.

Sosis, R. 2001, “Costly signaling and torch fishing on Ifaluk atoll”, *Evolution and Human Behavior*, 21, 223–244.

Spence, M. 1973, “Job Marketing Signaling”, *Quarterly Journal of Economics*, 87, 355–374.

- Spence, M. 2002, "Signaling in Retrospect and the Informational Structure of Markets", *American Economic Review*, 92, 434-459.
- Sripada, C. 2016, "Self-Expression: A Deep Self Theory of Moral Responsibility", *Philosophical Studies*, 173(5): 1203–1232.
- Sripada, C. 2017, "Frankfurt's Unwilling and Willing Addicts", *Mind*, 126(503): 781–815.
- Stout, N. 2020, "On the Significance of Praise", *American Philosophical Quarterly*, 57(3), 215-226.
- Strawson, G. 1986, *Freedom and Belief*, Oxford: Oxford University Press.
- Strawson, G. 1994, "The Impossibility of Moral Responsibility", *Philosophical Studies*, 75(1–2): 5–24.
- Strawson, P. F. 1962 [2008], *Freedom and Resentment*. Reprinted in Strawson, P.F. 2008, *Freedom and Resentment and Other Essays*, New York: Routledge.
- Talbert, M. 2012, "Moral Competence, Moral Blame, and Protest", *The Journal of Ethics*, 16(1): 89–109.
- Talbert, M. 2013, "Unwitting Wrongdoers and the Role of Moral Disagreement in Blame", in Shoemaker 2013: 225–245.
- Talbert, M. 2014, "The Significance of Psychopathic Wrongdoing", in Schramme 2014: 275–300.
- Talbert, M. 2022, "Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/fall2022/entries/moral-responsibility/>>.
- Tollefsen, D. 2017, "Epistemic Reactive Attitudes", *American Philosophical Quarterly*, 54.

- Vargas, M. 2005, “The Trouble with Tracing”, *Midwest Studies in Philosophy*, 29: 269–291.
- Vargas, M. 2006, “On the Importance of History for Responsible Agency”, *Philosophical Studies*, 127(3): 351–382.
- Vargas, M. 2013, *Building Better Beings: A Theory of Moral Responsibility*, New York: Oxford University Press.
- Veblen, T. 1899, *The Theory of the Leisure Class: An Economic Study of Institutions*. London: Unwin Books; reprinted in New York: Dover Publications, 1994.
- Wallace, R.J. 1996, *Responsibility and the Moral Sentiments*, Cambridge, MA: Harvard University Press.
- Wallace, R.J. 2011, “Dispassionate Opprobrium: On Blame and the Reactive Sentiments”, in Wallace, Kumar, and Freeman 2011: 348–372.
- Watson, Gary, 1975, “Free Agency”, *The Journal of Philosophy*, 72(8): 205–220.
- Watson, G. 2001, “Reasons and Responsibility”, *Ethics*, 111(2): 374–394.
- Watson, G. 1987 [2004], “Responsibility and the Limits of Evil: Variations on a Strawsonian Theme”, in Schoeman 1987: 256–286. Reprinted in Watson 2004: 219–259.
- Watson, G. 1996 [2004], “Two Faces of Responsibility”, *Philosophical Topics*, 24(2): 227–248. Reprinted in Watson 2004: 260–88.
- Watson, G. 2004, *Agency and Answerability: Selected Essays*, New York: Oxford University Press.
- Watson, G. 2011, “The Trouble with Psychopaths”, in Wallace, Kumar, and Freeman 2011: 307–31.
- Wolf, S. 1987, “Sanity and the Metaphysics of Responsibility”, in Schoeman 1987: 46–62.
- Wolf, S. 1990, *Freedom Within Reason*, New York: Oxford University Press.

Wolf, S. 2011, “Blame, Italian Style”, in Wallace, Kumar, and Freeman 2011: 332–347.