

**Unlinking Moral Responsibility and Appraisability: Defending an
Asymmetric Relation**

**by
Shathy Akter**

**A Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
of the University of Manitoba
in partial fulfilment of the requirements of the degree of
Master of Arts
Department of Philosophy
University of Manitoba
Winnipeg**

Copyright © 2026 by Shathy Akter

Abstract

In ordinary moral judgment, we often ask whether an action is related to an agent in the morally relevant sense: is the action really the agent's, and is the agent praiseworthy or blameworthy for doing it? These questions come together and make it natural to assume that the statuses of being responsible and being appraisable always travel together. I argue that these two are connected but they do not stand in a biconditional relation according to which an agent is morally responsible if and only if she is morally appraisable. Instead, I argue for an asymmetric relation between them: moral responsibility is necessary for appraisability, but it is not sufficient for it. The key to this asymmetric relation lies in distinguishing the agent's responsibility-relevant relation to the action from the further evaluative question of praise and blame. I develop this argument in three steps. First, I examine two influential families of theories in the responsibility literature to show how and why their views make the biconditional assumption attractive. Second, I challenge the sufficiency direction of the biconditional assumption through cases of moral luck and structural injustice. These cases are used to argue that an agent can be morally responsible for doing an action even though she is not praiseworthy or blameworthy for it. Finally, I provide a constructive account of responsibility and appraisability by explaining the conditions required for being responsible and appraisable. On my account, moral responsibility concerns whether an action is attributable to an agent as an exercise of her agency, and moral appraisability concerns whether the agent is praiseworthy or blameworthy for doing that action.

This asymmetric relation provides a clearer way to understand agency, praise, blame, luck, and constraint. Cases involving structural injustice and moral luck often create a moral dilemma in which we have to move either toward blame or toward the denial of responsibility. The asymmetric relation defended in this thesis helps avoid this moral dilemma and makes it possible to state a morally meaningful relation between agents and actions without immediately moving to praise or blame.

Acknowledgments

I would like to express my gratitude to my supervisor, **Professor Eliza Wells**. Her meticulous observations throughout the work helped me develop what I hope is a thought-provoking thesis. Every line of this thesis is a clear result of my supervisor's thoughtful feedback, which helped me structure the arguments of this thesis more clearly without losing my originality. I am grateful to **Professor R.J. Leland** for his constant guidance and invaluable feedback. He gave careful attention to every detail in the drafts of each chapter. His kindness and encouragement were among the reasons I could complete this thesis. He knew how to keep my morale high and created the environment in which I could present ideas and thoughts without hesitation, regardless of how they initially looked. I will forever be grateful to both of them for giving me the opportunity to work with them. I am also grateful to **Professor Mark Hudson** for having agreed to be the external member of my thesis committee and to **Professor Sarah Hannan** for chairing my thesis defense.

Professor Adam Murray had a huge role in this project, even though he was not directly involved in this project. He helped me when I needed it the most. At a crucial moment, he approached me not only as a teacher but also as a compassionate mentor. He advised me on how I could continue moving forward. I regret that I did not have more chances to meet **Professor Chris Tillman** during my coursework. Nevertheless, I experienced his affection before joining the program, even when I was in Bangladesh. He showed patience in answering my inquiries no matter how minor they were, and gave me a lot of encouragement to join the Philosophy program at the University of Manitoba. I am glad that I followed his encouragement and came here. My sincere thanks go to **Annette Jones** for her support. She showed care and concern for me, which reminded me of the love and care I received from my loved ones. I also thank **Steven Brown** for helping me with administrative matters. I thank my friend **Jenna Ibrahim** for her constant encouragement. I am grateful to **Gerry Beaulieu** for giving me the chance to work with him. The entire department has been supportive and kind to me. I will always keep those memories very close to my heart.

I owe my deepest gratitude to those who provided me with the constant emotional, spiritual, and material support that never let me give up in difficult times. My lifelong mentor, **Professor**

Mostofa Nazmul Mansur, is one of them. I never imagined that I could have a teacher like him who not only teaches but also helps shape thoughts in a meaningful way. Without his guidance, vision, and inspiration, nothing meaningful would have been achieved. Whatever success I may achieve in life, I will remain indebted to him. My gratitude extends to **Tasnim Irfan Keya** for keeping me motivated. I thank **Sanjida Tanjim** for all her help during this time. I am also grateful to **Tahomina Tanu** for our conversations during this crucial time. My profound thanks go to **Shamima Akter** for never letting me feel far from my loved ones. Life in Canada would have been difficult without her. She treated me as a member of her family and kept me hopeful. My thanks extend to my lovely roommates, **Puja** and **Akhi**.

The support I have always received from my family members is indescribable. Being raised by a single mother never made me feel helpless or hopeless because of the hero of my life, my mother, **Rashida Begum**. The way she wanted me to be perceived and the life she wished me to have cost her much of her own happiness. My brother, **Rafiqueal Islam**, is the person who protected me and taught me to dream big. My long journey from Bangladesh to Canada was made possible by his support. I am grateful to my sister-in-law, **Shahida**, for taking good care of my mother during my absence. My two little princesses, my nieces **Manha** and **Rayta**, have also been a constant motivation, as I know that, in their eyes, I am their hero who never gives up. Above all, I wish to express my utmost gratitude to the person who was there for me every step of the way, my husband, **Maksudul Islam**. He encouraged me to follow my dreams even though it distanced us by thousands of miles. During that time, he kept me away from stress so that I could focus only on my studies, even while he himself endured a major operation and loneliness. Life has blessed me with the great fortune of having a partner like him, who never ceased to offer patience, love, and support through all the ups and downs. This thesis would not have been possible without his constant support.

Dedication

In loving memory of

my father, **Sekandar Ali** (1965–2004)

and my grandmother, **Fuljan Bibi** (1930–2020)

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Dedication.....	v
 Chapter One—Introduction.....	 1
1.1 Moral Responsibility, Moral Appraisability, and the Biconditional Assumption.....	1
1.2 The Basic Concepts of Moral Responsibility and Moral Appraisability.....	3
1.3 Morally Complicated Cases: Motivating an Asymmetric Relation.....	7
1.4 Thesis Outline.....	13
 Chapter Two—Moral Responsibility and Moral Appraisability: the Biconditional Assumption.....	 14
2.1 Introduction.....	14
2.2 The Desert-Entailing Route.....	16
2.3 The Quality-of-Will Route.....	22
2.4 Two Routes to the Same Assumption: Differences and Common Ground.....	26
2.5 Conclusion.....	29
 Chapter Three—Responsibility without Appraisability: Challenging the Biconditional Assumption.....	 31
3.1 Introduction.....	31
3.2 Clarifying Responsibility and Appraisability.....	32
3.3 Responsibility without Appraisability: Moral Cases.....	36
3.3.1 Daniel: Moral Luck.....	36
3.3.2 Sara: Structural Injustice.....	41
3.4 What the Cases Show.....	44
3.5 Conclusion.....	45

Chapter Four—A Constructive Account of Moral Responsibility and Moral Appraisability	46
4.1 Introduction.....	46
4.2 Moral Responsibility: a Distinct Form of Agential Connection.....	47
4.2.1 Attributability: the Agent-Action Relation Required for Moral Responsibility.....	49
4.2.2 Required Control.....	50
4.2.3 Relevant Awareness.....	54
4.2.4 The Final Structure of Moral Responsibility.....	62
4.3 Moral Appraisability: Appraising the Agent for an Action.....	63
4.3.1 Evaluative Attitudes: the First Condition of Moral Appraisability.....	64
4.3.2 Manifestation: the Second Condition of Moral Appraisability.....	68
4.4 Completing the Asymmetry.....	70
4.5 Conclusion.....	73
Chapter Five—Conclusion	74
5.1 The Argument So Far and Its Scope.....	74
5.2 Looking Back at the Desert-Entailing and the Quality-of-Will Views.....	76
5.3 Responsibility without Appraisability: the Significance of the Asymmetric Relation.....	78
5.4 Implications for Morally Complicated Cases.....	80
5.5 Concluding Remarks.....	86
Bibliography	87

Chapter One

Introduction

1.1 Moral Responsibility, Moral Appraisability, and the Biconditional Assumption

In ordinary moral judgment, we often assume that being morally responsible for an action and being morally appraisable for doing it are closely connected. In this thesis, I understand moral responsibility as concerning whether an action counts as the agent's own exercise of agency. I understand moral appraisability as concerning whether an agent is praiseworthy or blameworthy for doing an action. These ideas are clarified in the next section and developed further in **Chapter Four**. For now, with this preliminary understanding in place, we can imagine a case that helps illustrate the ordinary connection between them:

Solo Suzy: Suzy is walking down the street. When she reaches the big house on the corner, she stops and considers. She has an intense dislike for the elderly couple who lives in the house, and she has just had an idea: she is going to upset them by breaking their window on the first floor. She carefully selects a stone and hurls it toward the window. She feels a jolt of satisfaction when she hears the sound of breaking glass. Then she walks on as if nothing has happened.¹

In the Suzy case, most of us will presumably think that she is responsible as well as blameworthy for doing the action. She intentionally throws the stone to disturb the elderly couple. When we ask if Suzy is blameworthy, we also ask whether she is responsible; if we think she is responsible for this harmful action, the natural conclusion is that she is also blameworthy.

¹Mattias Gunnemyr and Caroline Torpe Touborg, "You Just Didn't Care Enough: Quality of Will, Causation, and Blameworthiness for Actions, Omissions, and Outcomes," *Journal of Ethics and Social Philosophy* 24, no. 1 (2023): 1, <https://doi.org/10.26556/jesp.v24i.1696>.

This ordinary judgment about Suzy suggests that the questions of whether the action is the agent's doing and whether the agent is praiseworthy or blameworthy for that action arise together. This thought is also reflected in the influential approaches to moral responsibility in the recent philosophical literature. Some approaches explain responsibility in relation to deserved praise or blame, while others explain the relation through what an action reveals about the agent's will, attitude, concern, regard or disregard.² As I discuss in **Chapter Two**, both approaches take moral responsibility and moral appraisability as naturally connected—thereby motivating the thought I have presented through the Suzy case. I will show that they often rely on a **Biconditional Assumption** concerning the relation between responsibility and appraisability. According to this assumption, an agent is morally responsible for an action X if and only if she is morally appraisable for doing X.

With a simple case like Suzy's, we can see that responsibility and blameworthiness stand together. However, before analyzing how this relation stands in morally complicated cases, and how morally complicated cases may put pressure on the **Biconditional Assumption**, I will first clarify the basic concepts of moral responsibility and moral appraisability.

² For moral responsibility and desert see Derk Pereboom, *Free Will, Agency, and Meaning in Life* (Oxford University Press, 2014), 2; Dana Kay Nelkin, "Accountability and Desert," *The Journal of Ethics* 20, nos. 1–3 (2016): 173–89, <https://doi.org/10.1007/s10892-016-9230-0>; and Matt King, "Two Faces of Desert," *Philosophical Studies* 169, no. 3 (2014): 401–24, <https://doi.org/10.1007/s11098-013-0188-5>. For agents' will, attitudes, concerns see P. F. Strawson, "Freedom and Resentment," *Proceedings of the British Academy* 48 (1962): 187–211; R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge, MA: Harvard University Press, 1994); David Shoemaker, "Qualities of Will," *Social Philosophy and Policy* 30, nos. 1–2 (2013): 95–120.

1.2 The Basic Concepts of Moral Responsibility and Moral Appraisability

By moral responsibility, I understand the relation between an agent and an action. An agent is morally responsible for an action when the action is an exercise of her agency rather than something merely happening through her. In this sense, moral responsibility is stronger than causal responsibility. An agent may causally contribute to an event without being morally responsible for it. Suppose someone is standing in a crowded room and is suddenly pushed from behind; as a result, she breaks a flower vase next to her. Here, the agent is causally connected to the breaking of the vase as the flower vase breaks through the movement of her body. But the breaking of the flower vase is not connected to the agent as her action because her movement was caused by a sudden push from another person rather than by an exercise of her agency. Being morally responsible requires a more substantive agential connection between the agent and the action. We can, again, imagine this case in a modified manner that shows this required agential connection. Suppose that instead of being pushed from behind, the person becomes angry and deliberately pushes a table, knowing that the flower vase is likely to fall and break. In this case, the breaking of the vase is not merely something that happens through her body. It is connected to what she knowingly and intentionally does. That is, her action is connected to her as an exercise of her agency, and for that reason she is morally responsible for breaking the vase. Moral responsibility, as I use the term, requires this stronger connection between agents and actions. In **Chapter Four**, I will provide a detailed formulation of moral responsibility and its conditions.

Having explained the preliminary concept of responsibility, I now turn to moral appraisability. Like moral responsibility, moral appraisability also concerns a relation between agents and actions. However, it asks a different question. While moral responsibility asks whether an action counts as an agent's own exercise of her agency, moral appraisability asks whether the

agent is praiseworthy or blameworthy for that action. To explain this use of moral appraisability, I think it is necessary to clarify how I understand the terms “praise” and “blame” in my thesis. Praise and blame can be understood in more than one way. First, they can refer to the actual reactive attitudes, such as resentment, indignation, gratitude, or admiration. In this sense, praise and blame are used as attitudes that people have toward an agent. But the fact that someone feels resentment or admiration does not by itself show that the attitude is appropriate or the agent is praiseworthy or blameworthy. A community may resent or admire an action for morally distorted reasons. Haji’s “Forbidden Love” case illustrates this point. Haji describes how members of a rigid society may feel resentment or indignation toward agents involved in a socially transgressive relationship.³ But such attitudes, as we understand, are themselves morally objectionable. Hence, this shows that the attitudes people actually have toward an agent cannot by themselves determine whether or not an agent is morally praiseworthy or blameworthy.

An alternative interpretation is concerned with whether praise and blame would be appropriate, apt, or deserved. This is the sense of praise and blame (appraisal) I will be using throughout this thesis. To say that an agent is morally appraisable for doing an action *X* is not to say that people in fact praise or blame the agent for performing *X*. It is to say that the agent is praiseworthy or blameworthy for what she has done because of the way she is connected to the action, and the circumstances in which she performed the action. On this reading, appraisability is an evaluative notion. For the purpose of my thesis, I understand praise and blame as evaluative judgments that are directed at agents for doing a particular action. These judgements are concerned

³ Ishtiyaque Haji, *Moral Appraisability: Puzzles, Proposals, and Perplexities* (New York: Oxford University Press, 1998), 198.

with what an action reveals about an agent's exercise of agency in performing the action in question. I also think when praise and blame are understood in this sense, they are not essentially emotional. Reactive attitudes such as indignation, resentment, or admiration may accompany these judgments, but these attitudes are not required for the evaluative judgments.

With this understanding of praise and blame, I use moral appraisability in its narrower sense. According to this narrower sense, an agent is morally appraisable for a particular action when she is praiseworthy or blameworthy for it. The reason for this focus is that praise and blame are the forms of moral appraisal most directly and significantly connected to the discussion of responsibility mentioned in this thesis. As I will discuss in the next chapter, some views often connect moral responsibility with deserved praise or blame, while others connect it with the reactive attitudes through which agents are assessed as morally praiseworthy or blameworthy. This narrower focus also allows me to distinguish moral appraisability from moral assessment in its broader sense. Otherwise, one might read my thesis as a total denial of moral assessment whenever I say that an agent is neither praiseworthy nor blameworthy. I think that an action may be morally assessable in a broader sense even when it does not make the agent praiseworthy or blameworthy. We may assess an action as neutral, unfortunate, or morally burdensome without judging the agent praiseworthy or blameworthy. For example, an agent may perform an ordinary responsible action that reveals acceptable evaluative attitudes, but nothing morally significant for praise or blame. In such a case, a neutral assessment may be appropriate. And if moral appraisability is understood more broadly to include every kind of assessment, then an agent may be appraisable in this broader sense without being praiseworthy or blameworthy. However, this is not the sense I have used in my thesis while explaining appraisability. My use of appraisability is concerned with whether an agent is praiseworthy or blameworthy. So, the ordinary case in which the assessment is neutral is

different from cases in which praise and blame are appropriate in my appraisability sense, and from cases in which praise or blame is inappropriate because of luck. This difference will be visible in the next chapters.

On my view, appraisability does not follow simply from the fact that the action counts as the agent's own exercise of agency. Appraisability requires something further: the action must reveal the agent's evaluative attitudes and how they are expressed and enacted in deliberation, choice and action. By *evaluative attitudes*, I mean the agent's morally relevant attitudes toward the action or the situation: what the agent cares about, what she takes to matter, and how she is disposed to respond. These attitudes include explicit or implicit moral beliefs as well as concern, valuing, and attention to moral reasons. **Chapter Four** contains the fuller explanation of evaluative attitudes to explain how such attitudes need to be expressed and enacted, and thereby manifested, in deliberation, choice, and action for praise or blame to be fitting.

Now, to illustrate the idea of appraisability, we can consider an example. Suppose an agent named Emily is on her way to an important meeting and sees someone being harassed. She pauses and recognizes that the situation matters morally. If Emily stops to help, she would be both responsible and praiseworthy because it manifests concern, attention, and responsiveness to another person's need. By contrast, if Jenny, a bystander, recognizes the same situation but simply ignores it out of indifference and does nothing to help the person, then Jenny is blameworthy. The difference is not merely that one helps and the other does not. The difference also lies in what each agent's conduct reveals about the agent's evaluative standpoint and how these attitudes are enacted in deliberation, choice, and action.

Once we clarify the concepts of responsibility and appraisability, we see that these two notions are closely related. However, I will argue that they do not stand in a biconditional relation. More specifically, moral responsibility is necessary for moral appraisability, but it is not sufficient for it. The relation that I will defend is therefore asymmetric rather than biconditional. In the next section, I will present a morally complicated case to show how the ***Biconditional Assumption*** comes under pressure and how a situation like this motivates the asymmetric relation.

1.3 Morally Complicated Cases: Motivating an Asymmetric Relation

Our intuition concerning the relation between responsibility and appraisability seems consistent with ordinary moral judgment until we encounter some morally complicated cases. In such cases, the connection between being responsible and being appraisable is not always as direct as the ***Biconditional Assumption*** suggests. Agents may be involved in the same broad wrongdoing while differing significantly in motivation, fear, understanding, pressure, and practical options. These differences make it difficult to assume, without further argument, that responsibility and appraisability are always linked in the way the biconditional states. A historical illustration may help bring out this difficulty.

In 1994, over a period of roughly one hundred days, Rwanda witnessed one of the most notorious modern genocides. During these one hundred days, nearly one million people were killed—primarily members of the Tutsi ethnicity, but also members of other ethnicities in Rwanda who did not support the genocide. Although the killing was initiated by the Hutu militias, the violence was continued by ordinary citizens. The perpetrators were not people from outside of the communities where genocidal violence was carried out. Instead, they were people who were, only a week earlier, living alongside those they later killed. Neighbors, shopkeepers, farmers, teachers

and local administrators participated in the violence. Accounts from the survivors and researchers describe a striking variety of circumstances. Some people joined the violence enthusiastically. The motivation behind their killing was their aggression and hatred for the other group. Some people were driven by ideological zeal. They believed that Tutsis were dangerous enemies and that killing was necessary to protect their families or fulfill their civic duty. Others acted reluctantly under threats from local authorities or militia groups. Some even tried to avoid participating in the genocide for moral reasons but eventually gave in out of fear for their own safety or for the safety of their families. However, while there certainly were threats, propaganda, and fear, research indicates that the killings were often carried out in small local groups in which refusal was possible, and many individuals did refuse without being killed. Some Hutu civilians hid and protected Tutsis. Their act of saving people during that time demonstrated that non-participation, though risky, was within the realm of action for some.⁴ That means that, for some of them, having the chance to act differently also meant that they could act in ways that morality requires. They could refuse to participate in the killing or could help people when possible.

However, introducing this case here is neither meant to settle the moral status of every participant nor meant to suggest that their participation in this genocide is morally ordinary. The motivation instead is to bring out a morally complicated case where moving from responsibility to appraisability is not as easy as in an ordinary case like Suzy's. Since in this case participants act under different motivations, ideology, understanding, pressure and fear, the relation between responsibility and appraisability becomes difficult to assess. This case, thus, raises the central questions of this thesis: what is the relation between being morally responsible and being

⁴ Mahmood Mamdani, *When Victims Become Killers: Colonialism, Nativism, and the Genocide in Rwanda* (Princeton, NJ: Princeton University Press, 2001), 3–7.

appraisable for doing an action? Are they mutually entailing as the ***Biconditional Assumption*** suggests? Or, can an agent be responsible for an action even though she is not appraisable for doing it?

On first impression, most of us may readily judge those agents involved in killing as both responsible and blameworthy for doing these heinous actions. They had moral obligations to follow, which forbade violent participation in genocidal violence. They were able to understand what they were doing, and in many cases their agency was not wholly bypassed.⁵ However, once we consider the pressure, fear, motivation, and understanding under which different participants acted, the initial judgment becomes less straightforward. Someone who participated enthusiastically out of hatred is both morally responsible and blameworthy. By contrast, someone who acted reluctantly under serious threat raises a further question: does being morally responsible for participating in the wrongdoing always mean being praiseworthy or blameworthy for it, and can an agent be responsible without being appraisable? The simple reason for delving into these worries is that we think appraisal should not be assessed based only on what an agent does. A person may be responsible in the sense that the action is done by her own exercise of agency rather than causally connected to her. However, the fear and pressure under which the agent acts affect whether praise or blame is fitting. Here, an important distinction needs to be made to clarify the matter. One question is whether an agent is appraisable at all in the narrower sense used in this thesis. Another question is what kind or degree of appraisal is fitting. These questions should be

⁵ Not controlled by any mind-controlling intervener, such as a Frankfurt-style controller, or not under any influence that totally blinds the agent about the matter. See Harry G. Frankfurt, “Alternative Possibilities and Moral Responsibility,” *The Journal of Philosophy* 66, no. 23 (1969): 835–36, <https://doi.org/10.2307/2023833>.

kept separated because we usually think that some people are more blameworthy than the others instead of thinking that they cannot be morally assessed, and some people are responsible but they are not praiseworthy or blameworthy. In the Rwanda case, readers may be more willing to say that different participants are blameworthy to different degrees than to say that participants are not morally assessable. So, my claim in this thesis does not contradict the claim that agents can be morally assessed in a broader sense. Rather, it is focused on the question of whether being morally responsible always makes an agent praiseworthy or blameworthy.

Hence, the Rwanda case portrays how people involved in the killing differ in belief, motivation, and moral understanding, even when they perform the same morally horrific action of participating in genocide. These differences matter because appraisability concerns not only whether the action is the agent's own but also whether the action expresses and manifests the agent's evaluative attitudes. This is why morally complicated cases put pressure on the ***Biconditional Assumption***.

A defender of the ***Biconditional Assumption*** might respond that different participants did not really perform the same action in the morally relevant sense. For example, when some participants were involved in killing with malice and some acted reluctantly under threat, they should be understood as doing different actions under different descriptions. Even if we agree that such differences matter, we still need to ask whether the action is an exercise of the participants' own agency and whether or not that action makes them blameworthy. Knowing about the different scenarios of the Rwanda case, it is clear that the question of responsibility and the question of appraisability should not be assumed to have the same answer.

These distinctive questions of moral responsibility and moral appraisability show how the Rwanda case creates a problem for the ***Biconditional Assumption***. In ordinary cases, we may easily say that responsibility and appraisability share a biconditional relation as we saw in the Suzy case. But situations illustrated by the Rwanda case put pressure on that assumption because different participants in the genocide seem to raise different questions. We have learnt that some agents participated with hatred, and some with ideological zeal. People in these groups are morally responsible since their actions are connected to them; they are also morally appraisable because their actions reveal their attitudes and manifest their attitudes. However, other agents acted under fear, social pressure, or limited practical options. We agree that fear and pressure can sometimes undermine agency in important ways, and in some cases they can undermine blame. However, they do not always erase agency. Under fear and pressure, agency remains when an agent understands what she is doing, has some practical control and acts in ways that are not merely causally connected to her. For this reason, the action counts as an exercise of the agent's agency. But her actions do not necessarily manifest the kind of attitudes that would make the agent blameworthy because fear, pressure, and limited options under which she acts may affect whether she is blameworthy. This is where the ***Biconditional Assumption*** comes under pressure. If moral responsibility and moral appraisability are understood as traveling together, then once we say that an agent is responsible, we also have to say that the agent is praiseworthy or blameworthy. But if we apply that assumption to the Rwanda case, we fail to account for the fear, social pressure, and limited options under which many agents participated in the genocide. By contrast, if we say that they are not blameworthy or not blameworthy in the same way, then the ***Biconditional Assumption*** suggests that we deny that they are morally responsible. Or, we may have to reduce their relation to the wrongdoing to mere causal involvement. Either reading of this case thus leads to an

unconvincing result. The Rwanda case therefore shows that the relation between responsibility and appraisability cannot always be biconditional. Rather, how the relation should be understood needs careful examination.

This still might seem implausible to those who assume a tight conceptual link between responsibility and appraisability. However, I am not denying that they can be connected. Instead, my claim is that they are not linked by the biconditional relation. Being morally responsible for an action does not by itself make an agent praiseworthy or blameworthy for that action. However, I preserve the opposite direction: to be morally appraisable for an action, the agent must first be morally responsible for it. Praise and blame are fitting only when the action is connected to the agent in the relevant sense. The relation I defend is therefore asymmetric: moral responsibility is necessary for moral appraisability, but it is not sufficient.

The central aim of my thesis is to explain and defend this asymmetry. I will argue that the key to this asymmetry lies in distinguishing the agent's responsibility-relevant relation to the action from the further evaluative question of praise and blame. To defend this claim, I will distinguish the responsibility question of whether an action is connected to the agent as an exercise of her agency from the appraisability question of whether that action reveals the agent's evaluative attitudes in a way that makes her praiseworthy or blameworthy. Understanding this distinction will allow us to preserve the thought that an action can genuinely belong to an agent's agency, while also leaving open whether that action reveals the evaluative attitudes required for praise or blame. This distinction between responsibility and appraisability also has broader implications for morally complicated cases like *moral luck* and *structural injustice*, which I am going to discuss in later chapters.

1.4 Thesis Outline

This thesis proceeds as follows. In *Chapter Two*, I will examine why the ***Biconditional Assumption*** appears attractive in the literature on moral responsibility, and how different views motivate the ***Biconditional Assumption***. I will focus specifically on two routes: the ***Desert-Entailing*** route and the ***Quality-of-Will*** route. Although these routes differ in what they treat as explanatorily central, both make responsibility and appraisability appear to stand together.

In *Chapter Three*, I will then challenge the ***Biconditional Assumption***, particularly the sufficiency direction of the biconditional, by arguing that moral responsibility is not sufficient for moral appraisability. Through cases involving moral luck and structural injustice, I will show that an agent may be morally responsible for an action even when praise or blame is not fitting.

In *Chapter Four*, I will develop an alternative positive account of moral responsibility and moral appraisability. I will explain what conditions are required for being responsible and appraisable. This account explains the asymmetry defended in this thesis: moral responsibility is necessary for moral appraisability, but it is not sufficient.

Finally, I will end this thesis by exploring the broader significance of embracing the asymmetric relation between responsibility and appraisability. Additionally, I will show how morally complicated cases can have an alternative yet better explanation when they are observed through the asymmetric relation.

—

Chapter Two

Moral Responsibility and Moral Appraisability: the Biconditional Assumption

2.1 Introduction

Suppose Anna learns that her colleague Joana is being considered for a promotion. Out of jealousy, Anna spreads a false rumor that Joana cheated on an important project. Because of the rumor, Joana's reputation at her workplace is tarnished, and she loses her chance of getting the promotion. In this case, it seems natural and compatible with our everyday moral judgment that Anna is responsible as well as blameworthy for what she has done.

This kind of case highlights a familiar assumption in our moral thought. According to this way of thinking, responsibility and appraisability are not treated as separate questions. This means that if Anna is a morally responsible agent for spreading the rumor, she is also blameworthy for doing it; and when she is a blameworthy agent, it seems natural to assume that she must be responsible for the action. This assumption becomes philosophically significant because it is not only part of ordinary moral judgment but also appears in theories of moral responsibility that explain being responsible as closely connected to being appraisable. The ***Biconditional Assumption***, thus, arises from the above thought: an agent is morally responsible for an action *X* if and only if she is morally appraisable for doing *X*.¹ As mentioned earlier, here, moral appraisability is understood as praiseworthiness or blameworthiness.

¹ Angela Smith notes that versions of a similar biconditional appear in several contemporary accounts of moral responsibility, where responsibility is connected with eligibility for distinctive moral responses. She writes, "S is morally responsible for X if and only if S is an eligible candidate for certain distinctive moral responses on the basis of X." I use this point only to show that the connection between being morally

The goal of this chapter is not merely to explain why the assumption appears natural in the literature on moral responsibility, but to bring out how influential accounts often treat responsibility and appraisability in a way that makes them appear to travel together. At the same time, this does not mean that the supporters of the ***Biconditional Assumption*** confuse the concept of responsibility with appraisability. The problem, I will argue, is that their explanations of responsibility make it seem as though the two statuses always apply together: whenever an agent is responsible, she is appraisable, and whenever she is appraisable, she is responsible.

The supporters of the ***Biconditional Assumption*** have reasonable explanations for asserting this connection. I will focus on two influential families of theories about moral responsibility. The first route is concerned with deserved appraisal. On the ***Desert-Entailing route***, moral responsibility is understood as the kind of relation between an agent and an action that can make praise or blame deserved rather than unfair. What makes appraisal deserved, on this view, depends on whether the agent stands in the right responsibility-grounding relation to the action. The ***Desert-Entailing*** route explains responsibility through conditions such as control, understanding, fair opportunity, avoidability, ownership, or the agent's relation to the action. On this route, the conditions that make agents responsible are also treated as the conditions that make agents eligible for deserved appraisal.

responsible and being morally appraisable is a familiar assumption in the literature. However, I do not take moral appraisability to be identical with reactive attitudes or with the practice of holding responsible. Rather, the point is that many accounts explain being responsible in ways that bring it close to being an appropriate target of moral evaluation. This broader tendency is especially visible in quality-of-will views. Angela M. Smith, "Blame and Holding Responsible," in *The Oxford Handbook of Moral Responsibility*, ed. Dana Kay Nelkin and Derk Pereboom (New York: Oxford University Press, 2022), 270–71.

The second route is concerned with the will expressed in action. On the ***Quality-of-Will route***, responsibility is explained by what an action reveals about the agent's will, such as goodwill, ill will, concern, indifference, or disregard. An agent is responsible when the action is expressive of the attitudes that constitute her will. Since praise and blame are fitting responses to good or defective qualities of will, this route also makes responsibility and appraisability appear closely connected.

The reason for examining these two routes is to show how the ***Biconditional Assumption*** can be motivated in different ways. These two routes are not necessarily incompatible, but they differ in what they treat as explanatorily central. The ***Desert-Entailing*** views foreground the importance of fairness. For our moral appraisals to treat their targets fairly, those responses should only occur when agents stand in the relation to an action that makes them deserving of praise or blame. By contrast, the ***Quality-of-Will*** views focus on the fact that our appraisals are aimed at the moral quality of an agent's will. On this view, praise and blame respond to what the action reveals about the agent's will. They do not make fairness or desert the central issue. While these views do not explicitly endorse the biconditional as I have stated it, I will show that they often rely on that assumption. Moreover, I will return to the Anna case throughout the chapter to show how these two routes explain the same responsibility–appraisability judgment in different ways. In the next chapter, I will argue against this assumption by showing that moral responsibility is not sufficient for moral appraisability.

2.2 The ***Desert-Entailing*** Route

According to the ***Desert-Entailing*** views, responsibility is connected to deserved appraisal. This claim does not mean that being morally responsible is identical with being morally appraisable. Rather, it asserts that to be responsible is to stand in the kind of relation to one's action that can

make praise or blame deserved. It should not be mistaken for a causal connection between agents and actions. Additionally, merely saying that agents deserve blame just in case there are some features of the agents or some relation that they stand in that makes blame appropriate would not help much. If desert is understood as merely appropriateness to praise or blame, then responsibility would just be a synonym for praiseworthiness or blameworthiness. Similarly, if desert means only fitting appraisal, then the ***Desert-Entailing route*** would not be distinct from the ***Quality-of-Will route***, as I discuss in **Section 2.3**, since a quality-of-will theorist can also say that blame is fitting when an action expresses ill will, disregard, or defective concern.

To illustrate, in the case above, Anna deliberately spreads a false rumor about Joana because she is jealous of Joana's success. If we say that Anna deserves blame because blame is appropriate in this case, the explanation does not show what is distinctive about the ***Desert-Entailing*** views. A quality-of-will theorist can say the same thing in a different vocabulary: blame is fitting because Anna's action expresses jealousy, resentment, ill will, or disregard. In this Anna case, then, both the desert-entailing theorist and the quality-of-will theorist reach the same verdict. Unless desert is understood and explained more substantially, it risks collapsing into the idea that desert is just a synonym for the aptness of praise or blame. On the ***Desert-Entailing*** views, desert must be understood with something more substantial than merely the appropriateness or fittingness of praise or blame. With this approach, the concern shifts from whether the action reveals something morally good or bad about the agent to the further question of whether the agent had the relevant control, opportunity, ownership, or relation to the action such that praise or blame would be deserved rather than unfair.

Different desert-entailing theorists try to explain the more substantial notion of desert differently. The supporters of the ***Desert-Entailing*** views, such as Pereboom, King, and Nelkin do

not give the same account of desert, but they share the thought that praise or blame is fitting only when the agent satisfies further desert-grounding conditions.² Desert is not understood as something different from fittingness, rather it is a more specific form of fittingness for deserved praise or blame. It asks whether praise or blame would be fair to direct at the agent, given that the agent had control, ownership, or fair opportunity that would make praise or blame deserved rather than unfair. For some theorists, desert is explained through control, ownership, or the action's being connected to the agent. For others, it is explained through ownership, responsiveness to moral reasons, or fair opportunity to act according to one's agency.³ Even though different theorists may conceive of the conditions for desert differently, they all think that desert depends on more than fittingness of praise or blame. Dana Kay Nelkin's idea of desert can be useful to explain the desert view. Nelkin uses the concept of *fair opportunity* to explain the substantial notion of desert. Nelkin frames her discussion in terms of responsibility in what she calls the accountability sense and desert as sanction.⁴ But, here, I use her view to clarify the ***Desert-Entailing*** idea that responsibility and deserved blame depend on an agent's fair opportunity to act rightly. According to Nelkin's view, responsibility in the accountability sense is connected to

² See Derk Pereboom, *Free Will, Agency, and Meaning in Life* (Oxford: Oxford University Press, 2014), 2; Matt King, "Moral Responsibility and Merit," *Journal of Ethics and Social Philosophy* 6, no. 2 (2012): 1–18; Matt King, "Two Faces of Desert," *Philosophical Studies* 169, no. 3(2014): 401–24, <https://doi.org/10.1007/s11098-013-0188-5>; Dana Kay Nelkin, "Desert, Fairness, and Resentment," *Philosophical Explorations* 16, no. 2 (2013): 117–132, <https://doi.org/10.1080/13869795.2013.787438>; Dana Kay Nelkin, "Accountability and Desert," *The Journal of Ethics* 20, nos. 1–3 (2016): 173–189, <https://doi.org/10.1007/s10892-016-9230-0>.

³ Pereboom, *Free Will, Agency, and Meaning in Life*, 2; Nelkin, "Accountability and Desert," 184–86; Dana Kay Nelkin, "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness," *Nous* 50, 2 (2016): 371–74, <https://doi.org/10.1111/nous.12079>; King, "Two Faces of Desert," 401–3.

⁴ Nelkin, "Accountability and Desert," 184–86.

demands: demands to act, or refrain from acting, for the right reasons. These demands are appropriate only when agents have a fair or reasonably high quality-opportunity to meet them.⁵ Nelkin argues that the same condition also applies to deserved blame: an agent deserves blame when she acts wrongly even though she has a fair opportunity to avoid wrongdoing.⁶ I take this to mean that an agent would be both responsible and blameworthy when she had the kind of opportunity needed for the action to be fairly connected to her agency. Otherwise, it would be unfair to judge her responsible and blameworthy for things that she had no real opportunity to avoid.

In my view, the fairness point is important because a responsibility judgment is fair only when it is made on the basis of reasons relevant to that kind of judgment. For desert theorists, the relevant reason is connected to an agent's own exercise of agency. If an agent lacked a fair opportunity to avoid wrongdoing, then judging her responsible and blameworthy would not be based on the right kind of reason. It would be like treating the agent as responsible for something that was not available to her as an exercise of agency. This is why lack of fair opportunity undermines moral responsibility. A manipulation case can be used to demonstrate the above point. Suppose an agent says something cruel because another person has covertly shaped her beliefs, motives, or practical reasoning in order to bring about that action. Since the agent's deliberation has been directed by another person's intervention, her actions are being used to satisfy the manipulator's purpose, rather than being exercised through her own normal control. In that sense, she lacks a fair opportunity to recognize the moral reason against acting cruelly or to guide herself differently. Here, manipulation prevents the agent's action from being connected to her in the right

⁵ Nelkin, "Accountability and Desert," 185.

⁶ Nelkin, "Accountability and Desert," 184-85.

way for deserved blame even though the action appears morally defective. For that reason, the desert-entailing theorist says that blame is not deserved in such cases.

Understanding desert in this more substantial sense allows us to state more clearly what it means to be morally responsible according to the ***Desert-Entailing*** views. An agent is morally responsible when she satisfies the conditions that make praise or blame deserved rather than unfair. These conditions may be explained in different ways, but they include control, understanding, fair opportunity, avoidability, ownership, and the agent's relation to the action. These conditions determine when the action is fairly attributed to the agent in a desert satisfying sense. If an agent lacked control, acted in unavoidable ignorance, or could not reasonably have been expected to avoid the action, then blame would be unfair, because the action would not be connected to the agent's agency in the way desert requires. By contrast, when these excusing conditions are absent, the agent's action is connected to her agency and praise or blame is deserved.

Thus, we can say that, on the ***Desert-Entailing*** route, the same conditions that establish responsibility are also treated as the conditions that make the agent eligible for deserved appraisal. When an agent has the relevant control, understanding, opportunity, and ownership conditions, the action is connected to her agency in the way required for responsibility. Since the action is connected to her in the right way, praise or blame would also be based on the right reason. The agent is therefore responsible and praise or blame would be deserved. If these conditions are absent, then the agent is not morally responsible in the relevant desert-entailing sense. And if the agent is not morally responsible, then praise or blame cannot be deserved. This is how the ***Desert-Entailing*** route makes responsibility and appraisability appear to stand or fall together.

With the above understanding of the desert view, we can move on to how cases of excuse can be explained. For that purpose, I will highlight another case parallel to Anna's, namely Ben's. Ben's case helps show how this desert-entailing structure works in an ordinary case of *excuse*. Suppose Ben repeats the same rumor about Joana, but he does so while mistakenly believing that it is true and without understanding its harmful significance. On the ***Desert-Entailing*** route, Ben's ignorance matters because it makes blame unfair: he lacks the relevant awareness needed to avoid the wrong in the way Anna possessed. And we don't expect people to act better or differently than what they do if they do not know better or different than what they know. Even though Ben repeats the rumor, he does not understand its falsity or harmful significance. His ignorance blocks the desert-grounding connection between the action and his agency. On the ***Desert-Entailing*** route, this same excusing condition also undermines responsibility: the condition that makes blame undeserved is also the condition that prevents Ben from being responsible in the relevant sense. Anna, by contrast, deliberately spreads the rumor, understands what she is doing, and has the relevant control and opportunity to avoid doing it. No excusing condition blocks the connection between her action and her agency. Anna is, therefore, both responsible and blameworthy.

One of the driving forces behind this view has been to explain why responsibility has a distinctive moral significance. Desert-Entailing theorists understand moral responsibility as involving the conditions under which moral appraisal would be deserved rather than unfair. They think that this feature makes moral responsibility distinct from causal or legal responsibility.⁷ Also by connecting responsibility and appraisability, this view ensures that responsibility remains meaningful. To sum up the general view of desert, responsibility is not first established

⁷ Matt King, "Moral Responsibility and Merit," *Journal of Ethics and Social Philosophy* 6, no. 2 (2012): 1.

independently, with appraisal added afterward. Rather, being responsible already places the agent within the space of moral appraisal, where she can be found worthy of praise or blame.

From the above discussion of the ***Desert-Entailing*** views, the argument supporting the ***Biconditional Assumption*** can be deduced. The argument goes as follows: if an agent is morally responsible for *X*, then she is morally appraisable for *X* because what makes her responsible is the same thing that makes praise or blame deserved. Conversely, if an agent is genuinely praiseworthy or blameworthy for *X*, then that appraisal must be deserved; and if it is deserved, it must rest on the features that make the agent responsible in the relevant sense.

In this way, the ***Desert-Entailing*** route provides one explanation of why responsibility and appraisability are often treated as standing together. At the same time, this view shows that the ***Biconditional Assumption*** has a real philosophical motivation: if responsibility is understood through desert, then it becomes difficult to separate being responsible from being eligible for praise or blame because separating them might make the idea of responsibility trivial. However, this connection is not limited to desert-based accounts. Versions of a similar assumption are also present, either explicitly or implicitly, in other parts of the responsibility literature. In the next section, I will explain how this assumption appears in the ***Quality-of-Will*** views.

2.3 The ***Quality-of-Will*** Route

According to the ***Quality-of-Will*** view, an agent is morally responsible when her action reveals something morally significant about her attitudes, intentions, concern, or disregard. The motivation behind this view is that we want to know what agents' actions reveal about their wills or character. An agent is morally responsible when the action is expressive of the attitudes that constitute her will. When an action reveals the agent's will in this way, praise or blame becomes fitting as a

response to the will expressed. This does not mean that only the quality-of-will theorists care about the agents' wills expressed in actions. The ***Desert-Entailing*** views may also agree that cruel intentions, indifference, or disregard help explain why blame is apt. But the distinctive claim of the ***Quality-of-Will*** route is that no further desert or fairness is required beyond the quality of will expressed in the action. If the action manifests goodwill, praise is fitting; if it expresses ill will or disregard, then blame is fitting. Ignorance or lack of control can make blame inappropriate, but they do so because they undermine the degree to which actions express ill will. Moral responsibility is, therefore, understood through the quality of will expressed in actions.

We see that this view gives a different way of explaining the responsibility-appraisability relation from the ***Desert-Entailing*** route. The ***Desert-Entailing*** route asks whether praise or blame would be fair or deserved, given the agent's control, opportunity, ownership, or relation to the action. The ***Quality-of-Will*** route does not need to explain appraisal in that way. The reason is that on the ***Quality-of-Will*** route, the same feature explains both responsibility and appraisability. If an action is expressive of the attitudes that constitute the agent's will, then the action is already connected to the agent in the responsibility-relevant way. And if the will expressed is morally good or defective, then praise or blame becomes fitting as a response to that expression. In that sense, once the quality-of-will theorists identify the action as expressing the agent's will, much of the work has already been done: the feature that makes an agent responsible is also the feature that makes the agent praiseworthy or blameworthy.

Recalling the Anna-Ben case shows how the difference works. We know that Anna deliberately spreads the rumor whereas Ben repeats the same rumor while mistakenly believing it to be true and not realizing that it would harm Joana's standing at work. Although both are involved in spreading the same rumor, their actions do not have the same moral significance from the

quality-of-will perspective. The explanatory work, for the quality-of-will theorist, is done by the fact that Anna's action expresses a defective quality of will, while Ben's does not. Desert-Entailing theorists may agree that this expression of will matters, but to them expression is not what does the main explanatory work. For desert-entailing theorists the work is done by checking whether the agent stands in the right responsibility grounding relation to the action.

Therefore, the **Quality-of-Will** route connects responsibility and appraisability through expression: the agent is responsible because the action she does expresses her will, and the agent is appraisable because the will expressed is morally good or defective. However, it does not mean that Ben can never be blameworthy for his ignorance according to the **Quality-of-Will** views. Ben can be blameworthy for his ignorance if his ignorance is understood as something as a result of his lack of responsiveness to the seriousness of the moral issue at stake. Even if Ben's action does not reveal hostility, it can also be an expression of unresponsiveness to things that matter morally, which is also a form of ill will. This point also highlights how cases of moral ignorance are dealt with in the **Quality-of-Will** views of moral responsibility. By contrast, the **Desert-Entailing** views may excuse Ben because blaming him would be unfair for things that he could not have known better.

The above contrast can also demonstrate the general structure of the **Quality-of-Will** view that tightly links moral responsibility with moral appraisability. On this view, to be morally responsible for an action is for the action to be expressive of the attitudes that constitute the agent's will. Responsibility is therefore not first established through a separate desert or fairness condition, with appraisal added afterward. Rather, responsibility is explained by the same feature that makes praise or blame apt: the quality of will expressed in action. In this way, the **Quality-of-Will** route makes responsibility appear both necessary and sufficient for appraisability. It appears necessary

because praise or blame would not be fitting unless the action expresses the agent's will in the relevant way. It appears sufficient because, once the action expresses a morally good or defective quality of will, the agent becomes an appropriate subject of praise or blame.

This line of thought appears in different forms within the quality-of-will tradition. Taylor Madigan helps make the structure explicit by asserting that if being responsible is explained by the appropriateness of praise and blame, and praise and blame are fitting because of the agent's quality of will, then being responsible and being appraisable appear to stand together.⁸ Strawson gives this route its classic interpersonal form. In "Freedom and Resentment," he shifts attention away from abstract metaphysical debates about determinism and toward the ordinary moral relationships in which we respond to one another. For Strawson, what matters in these responses is not merely the causal origin or consequence of an action, but the quality of will expressed toward us.⁹ This is why we may feel less gratitude when someone benefits us accidentally than when she helps us intentionally, even if the outcome is the same. What differentiates our attitude is the person's attitude and intention toward us.¹⁰ The **Quality-of-Will** route therefore supports the

⁸ Taylor Madigan, "Two Dimensions of Responsibility: Quality and Competence of Will," *Journal of the American Philosophical Association* 10, no. 2 (2024): 282, <https://doi.org/10.1017/apa.2023.2>.

⁹ P.F. Strawson, "Freedom and Resentment," *Proceedings of the British Academy* 48 (1962): 190–93; R. Jay Wallace develops the same route in a more explicitly normative form through his account of moral demands and quality of choice; see R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge, MA: Harvard University Press, 1994), 127–31.

¹⁰ However, mentioning Strawson's discussion here might invite confusion unless I clarify how I am interpreting his view in this project. Some versions of this route are framed in terms of reactive attitudes or responsibility-related responses, while my concern is with moral appraisability understood as being praiseworthy or blameworthy. I do not identify reactive attitudes themselves with appraisability. Rather, I use this tradition to show how responsibility can be explained through a feature that also matters for

Biconditional Assumption by treating responsibility and appraisability through the same expressive relation between an agent and an action. In the next section, I will focus on the differences and common ground between the ***Desert-Entailing*** route and the ***Quality-of-Will*** route.

2.4 Two Routes to the Same Assumption: Differences and Common Ground

The ***Desert-Entailing*** route and the ***Quality-of-Will*** route provide two different approaches to the biconditional relation between responsibility and appraisability. These routes do not explicitly aim to defend the ***Biconditional Assumption***. The assumption emerges from the way each route explains moral responsibility. In explaining the assumption, they have adopted different features as central. The ***Desert-Entailing*** route asks whether an agent stands in the kind of relation to the action that makes praise or blame deserved rather than unfair. The ***Quality-of-Will*** route, on the other hand, asks what the action expresses about the agent's will, attitudes, regard, or concern.

One of the similarities between the ***Desert-Entailing*** and the ***Quality-of-Will*** views is that both are concerned with the relation between agents and actions, even though they interpret that relation differently. In this sense, the “agency” emphasized by the ***Desert-Entailing*** views and the “will” emphasized by the ***Quality-of-Will*** views are closely related. Both concern the connection between the agent and what she does, but they use that connection for different explanatory purposes.

appraisal: the quality-of-will expressed in action. In this way, the quality-of-will route helps explain why responsibility is often understood as placing the agent within the space of moral evaluation.

The difference between them can be seen by considering how each route treats cases involving loss of control. A desert-entailing theorist cares about loss of control because it is unfair to blame someone if they did not have control. This is a moralized concern about whether agents are fairly appraised. A quality-of-will theorist may also care about loss of control, but for a different reason. On this route, loss of control matters because it may prevent the action from expressing the agent's quality of will. If the agent's will is not really expressed in the action, then the action may not reveal her attitudes, concern, or disregard.¹¹

The deeper contrast concerns what each route treats as central. For the quality-of-will theorist, if the action is expressive of the agent's will, then the central issue is largely settled. If the will expressed is morally good or defective, then praise or blame is fitting as a response to that expression. Desert-Entailing theorists, by contrast, may accept that the action appears to express a defective quality of will, but they do not consider this to explain their theories. They ask a different question: does the agent stand in the right relation to the action, through control, fair opportunity, avoidability, ownership, ownership, or responsibility for the will, such that blame would be deserved rather than unfair?

We can imagine the Anna case in a modified manner so that it can be used to show how the verdicts for the same case differ between the *Desert-Entailing* and the *Quality-of-Will* views. Suppose Anna's action still appears to express hostility or disregard. But her agency is shaped by some kind of manipulation or coercive pressure in such a way that she lacks an opportunity to avoid acting as she does. We can also imagine that the presence of manipulation limits Anna's opportunity to avoid spreading the rumor but it does not create or replace her hostile behavior

¹¹ I Thank Professor R.J. Leland for bringing this point of contrast to my attention.

toward Joana. Anna's action may therefore continue to express her ill will, even though she lacks a fair opportunity to avoid acting as she does. When the ***Desert-Entailing*** view evaluates this kind of case it asks whether the agent deserves praise or blame based on what she does, how she was when she did the action, what she knew and what control she had. In the modified Anna case, we see that Anna lacks any real opportunity to avoid acting as she does because of manipulation. Hence, there was nothing she could have done other than what the manipulator wanted her to do. In this case, it would not be fair to think that Anna deserves blame for what she did. In that case, the quality-of-will theorist may continue to focus on what the action expresses about Anna's will. If the action expresses ill will or disregard, then blame may appear fitting on the quality-of-will grounds. By contrast, we have seen above that the desert-entailing theorist asks a different question: did Anna have the control, avoidability, fair opportunity, or ownership needed for blame to be deserved rather than unfair? This is what makes the ***Desert-Entailing*** route distinct from the ***Quality-of-Will*** route. It is not concerned with adding an extra condition after an agent's quality of will is identified. Rather, it explains responsibility and deserved appraisal through conditions under which praise or blame would be deserved.

Even with the differences in verdict and in explanation, the two routes share an important common ground. Both explain moral responsibility in a way that places the agent within the space of moral appraisal because they think that moral responsibility would lose its significance otherwise. Without appraisability, the difference between a tornado destroying someone's house and an agent intentionally destroying it is blurred. This shows their concern for two important phenomena: whether an action is the agent's in the responsibility-required sense, and whether the agent is praiseworthy or blameworthy for doing it. This is why both routes support the ***Biconditional Assumption***. Additionally, this shared structure is important because it shows why

the ***Biconditional Assumption*** is worth examining. It arises from serious attempts to explain why agents like Anna are appropriate subjects of blame while agents like Ben are not.

2.5 Conclusion

This chapter has examined how the ***Biconditional Assumption*** concerning the relation between moral responsibility and moral appraisability is made attractive in the literature on moral responsibility. The two approaches discussed in this chapter explain the connection in different ways. For the ***Desert-Entailing*** view, the connection appears plausible because moral responsibility is not treated as mere causal involvement. Rather, responsibility is understood through the conditions under which praise or blame would be deserved rather than unfair. On this route, the features that make an agent responsible are also treated as the features that make moral appraisal deserved. For the ***Quality-of-Will*** view, the connection appears plausible in a different way. This view explains responsibility by asking what the action reveals about the agent's will, concern, regard, indifference, or disregard. If an action is expressive of the attitudes that constitute the agent's will, then the action is connected to the agent in the responsibility-relevant way. If the will expressed is morally good or defective, then praise or blame becomes fitting. On this route, therefore, responsibility and appraisability appear closely connected because both depend on the quality of will expressed in action. Their explanations make the statuses of being responsible and being praiseworthy or blameworthy stand together. They do not necessarily state the ***Biconditional Assumption*** in the way I have used it in this thesis. But they explain moral responsibility in a way that places the agent within the space of praiseworthiness or blameworthiness. This shared structure is what I have called the ***Biconditional Assumption***.

This shared assumption is what the next chapter challenges. The challenge, nevertheless, is not to deny that moral responsibility has moral significance. Rather, what I will challenge is the assumption that responsibility has such significance by standing in the biconditional relation. In the next chapter, I will argue that moral responsibility is not sufficient for moral appraisability. In some cases, an agent may remain morally responsible for an action *X* even though she is neither praiseworthy nor blameworthy. Showing this will make room for the constructive account in which I will give the conditions of moral responsibility and moral appraisability.

—

Chapter Three

Responsibility without Appraisability: Challenging the Biconditional

3.1 Introduction

In the previous chapter, I have examined two influential approaches to the relation between moral responsibility and moral appraisability: the *Desert-Entailing* views and the *Quality-of-Will* views. From these views, I have also drawn out a shared assumption. According to that assumption, responsibility and appraisability share a biconditional relation. The assumed relation can be stated as follows:

an agent is morally responsible for an action *X* if and only if she is morally appraisable for *X*.¹

This assumption plays a crucial role in theorizing responsibility because it captures an implicit way in which the significance of moral responsibility is often preserved in the literature. If responsibility were separated entirely from appraisability, one might worry that responsibility would become trivial. However, this does not mean that every account of moral responsibility takes this assumption in the way it is stated above. Rather, these two routes display an influential pattern in the literature that is responsibility is often explained in a way that makes it appear to stand or fall with appraisability.

In *Chapter One*, I have explained that praise and blame can be used in two senses. In one sense, praise and blame refer to actual responses such as resentment, indignation or admiration. Here, these responses are used as attitudes that people show toward others. The second sense understands praise and blame as evaluative judgments about whether agents are praiseworthy or blameworthy for doing actions. For the purpose of my thesis, I have adopted

¹ Here, moral appraisability means being morally praiseworthy or blameworthy for *X*, rather than being morally assessed in a broader sense.

the judgmental reading of praise and blame that concerns whether an agent is praiseworthy or blameworthy.

However, both senses of appraisability can pose problems when they are used in defining responsibility. If we define responsibility directly through the reactive attitudes, holding that an agent is responsible for an action whenever she is actually susceptible to attitudes such as resentment, indignation, gratitude, or admiration, then responsibility depends too much on contingent social practices rather than on the normatively grounded relation between agents and actions. By contrast, if responsibility is tied to judgmental appraisal, such that to be morally responsible is to be in a state such that praise or blame would be apt for what one has done, then responsibility loses its independent content. It becomes difficult to explain responsibility without appealing to appraisability. In that way, responsibility and appraisability are treated as conceptually inseparable from the beginning. But I think we have reasonable explanations to keep them separated. For there are cases in which agents are clearly responsible for their actions, yet it would be inappropriate or irrelevant to judge them praiseworthy or blameworthy. These cases will challenge the sufficiency direction of the ***Biconditional Assumption***: if an agent is morally responsible for *X*, then she is morally appraisable for *X*.

This chapter aims to show that these difficulties can be avoided, even though praise or blame are understood as evaluative notions, by separating responsibility from appraisability rather than treating them as conceptually inseparable. In the next section, I will clarify the working concepts of moral responsibility and moral appraisability for guiding the argument of this chapter.

3.2 Clarifying Responsibility and Appraisability

Based on the clarification in the previous section, it should be mentioned that I do not define responsibility in terms of praiseworthiness or blameworthiness. I will defend fully in ***Chapter Four*** that an agent is morally responsible for an action when she is related to that action in a

way that makes it attributable to her as an exercise of her agency. By “attributability” I mean an agent-action relation in which the action counts as the agent’s own action in a morally relevant sense, rather than as something happening through her. When this relation obtains, the action counts as her own in a morally relevant sense. I will return to the explanation of attributability and how it is obtained in the next chapter.

Moral appraisability, by contrast, concerns whether an agent is praiseworthy or blameworthy in virtue of what the action reveals about her evaluative attitudes. I will explain in **Chapter Four** that appraisability requires more than that the action is connected to the agent; it requires that the agent’s evaluative attitudes be expressed and enacted in action. To recall, by evaluative attitudes, I understand the attitudes through which an agent takes the morally significant features of an action or situation to matter. They include explicit or implicit beliefs, valuing, care, and attention to moral reasons. They may also be expressed through relevant dispositions, such as courage, conscientiousness, or, negligence. But this does not mean that every attitude or mood present in the agent is relevant to appraisability. Only those features that bear on how the agent responds to the morally significant aspects of the action or situation are relevant.

For appraisability, I argue that, it is not enough that the action is attributable to the agent in the way secured by moral responsibility. Appraisability also requires that the agent’s evaluative attitudes are expressed and enacted in deliberation, choice, and action. In analyzing moral appraisability, I will focus on a particular action rather than with a global assessment of the agent’s character. Character may be relevant in the background, but on my view appraisability is grounded more immediately in how the agent’s evaluative attitudes are expressed and enacted in deliberation, choice, and action in the case at hand. That means that the central question is not whether the agent has a generally good or bad character but whether

her exercise of agency in relation to *X* reveals something that makes her praiseworthy or blameworthy for *X*.

This account of moral appraisability also helps us clarify why appraisability should be distinguished both from responsibility and from the *Quality-of-Will* view. My aim is not to stipulate a different use of appraisability, but to capture the phenomenon of what makes an agent praiseworthy or blameworthy for a particular action that distinguishes it from responsibility. The *Quality-of-Will* view is about what an action expresses about an agent's will and tracks the moral quality of will. My account agrees that what an action reveals about an agent's attitudes matter, but it also asks how the agent's evaluative attitudes are expressed and enacted in action. So, my account of appraisability tracks whether an agent's evaluative attitudes are expressed as well as manifested in the action in a way that reveals something morally relevant about the agent's exercise of agency. To clarify this point, we can recall the Emily case from *Chapter One*. Emily, on her way, sees someone being harassed on the street. She decides to help the person. For helping the person, Emily is praiseworthy. A quality-of-will theorist will explain Emily's praiseworthiness by saying that her action expresses goodwill and concern for the person in distress. My account of appraisability agrees with this explanation, but it also asks how Emily's concern is manifested or how it is expressed and enacted in action. Since Emily notices that someone needs help, considers what the victim needs and chooses a way to help with care and attentiveness, her action reveals her goodwill as well as the appropriate manifestation of her evaluative attitudes in deliberation, choice, and action. Conversely, someone with the same intention as Emily's may act carelessly by ignoring risks, or by becoming involved in a way that shows negligence. Here, the mere presence of goodwill is not enough to make the agent praiseworthy. This is why I understand moral appraisability in a sense that is not only about whether an action expresses goodwill or ill will, but also about how the agent's evaluative attitudes are expressed and enacted in action.

These different focuses make responsibility and appraisability distinct from each other. Responsibility asks whether an action is properly connected to an agent as an exercise of her agency, whereas appraisability asks whether that action reveals and manifests evaluative attitudes that make the agent praiseworthy or blameworthy. This distinction allows us to challenge the assumed biconditional relation between responsibility and appraisability. This chapter particularly aims to challenge one direction of the relation: if an agent is responsible, then she is morally appraisable. That means that I think responsibility is not sufficient for appraisability. Conversely, I preserve the other direction of the biconditional: if an agent is morally appraisable for *X*, then she is morally responsible for *X*. That is, responsibility is necessary for appraisability. This is because appraisal is an evaluation of how an agent's evaluative attitudes are expressed and enacted in her action. But if the action is not attributable to the agent's agency in the right way, then it does not make sense to treat it as expressing anything morally significant or relevant about her attitudes. In that case, praise or blame would lack normative grounding. This does not only indicate that such appraisal would be unfair, but that it would be misdirected, since the agent would not stand in the right relation to the action for it to disclose what she takes to matter and how that is manifested through her action. Therefore, the result is an asymmetry: *responsibility is necessary but not sufficient for appraisability*.

The strategy of the remainder of this chapter is straightforward. If we can show a counterexample in which an agent is responsible for doing *X*, but she is not appraisable for *X*, then it would be enough to refute the claim that responsibility is sufficient for appraisability. But I think there are many such counterexamples. In what follows, I will present two cases that are designed to illustrate such counterexamples. These cases are not offered merely to elicit isolated intuitions. They also draw attention to the features that generate our intuitions, and so help explain how responsibility and appraisability can come apart. The other direction of the

relation, which is whether responsibility is necessary for appraisability, will be addressed in the next chapter. For our present purpose, the aim is more limited but decisive: to show that responsibility and appraisability can diverge, and that the assumed entailment is therefore mistaken.

3.3 Responsibility without Appraisability: Moral Cases

In this section, I present two moral cases. These cases aim to challenge the claim that moral responsibility is sufficient for moral appraisability. As mentioned earlier, if the biconditional were correct, then whenever an agent is morally responsible, she would thereby be an appropriate object of praise or blame. The two cases considered below are designed to demonstrate different routes by which moral responsibility and moral appraisability can come apart. The first case involves moral luck, and the second one involves structural injustice. Together, they show that there are cases in which agents appear responsible in a morally significant sense, yet praise or blame seems *irrelevant, inappropriate, or misplaced*.

I begin with the moral luck case in which our judgments of responsibility and appraisability seem quite baffling. Initially, the case tends to support the biconditional claim, but closer analysis highlights a different mechanism by which responsibility and appraisability diverge.

3.3.1 Daniel: Moral Luck

Suppose an agent, Daniel, decides to drive home during heavy rain. Although the rain required extra caution, the conditions were not so dangerous that driving itself was reckless or unreasonable. Like any driver, he has moral obligations to follow traffic rules, exercise reasonable care, and avoid imposing unnecessary or unreasonable risks on others. Daniel is a competent driver who usually follows whatever is required when driving. Moreover, his evaluative attitudes, understood as what he takes to matter morally in this situation and how he is disposed to respond, are, morally appropriate. He believes in careful driving and tries to manifest that belief in his conduct. However, when he approaches a four-

lane intersection, a child suddenly runs into the road from between parked cars. The child appears so abruptly that Daniel has almost no time to think over the matter. Daniel immediately applies the brakes, but the situation unfolds too quickly for him to avoid hitting the child.

Since the aim of the case is to show whether responsibility and appraisability can come apart with respect to the same object, we can imagine that the action, *X*, denotes Daniel's causing harm to the child through his risk-imposing act of driving. At first glance, the case might invite us to the following reasoning: Daniel was under an obligation to avoid harming others while driving. Since harm occurred, it might seem that the obligation was not fulfilled, and that Daniel therefore violated his obligation. And if he violated his obligation, then he would be blameworthy for the injury. This reasoning seems appealing because it moves directly from the harmful outcome to obligation violation, and from there to blame, in a way that appears consistent with the biconditional claim. However, I think this reasoning rests on a mistaken assumption. The assumption is that the mere occurrence of harm shows that the agent failed in a way that reflects moral fault. This does not mean that Daniel is not responsible for *X*. Here, the responsibility-relevant point is that Daniel is not a bystander or someone merely connected to the event, but he is someone who can understand his action and guide his action accordingly, and the child's harm occurs through his risk-imposing action.

At the same time, the case should not be understood as one in which Daniel is blameworthy for *X*. Daniel does not ignore the risk, act recklessly, or show indifference to the child's safety. Rather, he tries to exercise agential control so as to prevent the risk of harm from eventuating. For that reason, blame for *X* seems misplaced. If the ***Biconditional Assumption*** were true, this absence of blameworthiness would also suggest that Daniel is not responsible for *X*. However, this is precisely the conclusion I want to resist. On my view, Daniel is morally responsible for *X* because *X* is attributable to him as an exercise of his agency. He knowingly undertakes an activity that imposes a conceivable risk on others. He understands what he is

doing and guides his action accordingly. These features are part of my positive conception of responsibility which I will defend in the next chapter. His responsibility, then, is not merely causal. It lies in the fact that his agency is linked to the child's harm through a knowingly risk-imposing action. Daniel did his best to exercise agential control so as to prevent the risk of harm from eventuating. However, he got behind the wheel knowing that driving in such conditions imposes some level of risk on third parties, despite his conscientious driving behavior. And this makes him morally responsible for X. But this responsibility should not be taken to imply blameworthiness; that is precisely the sufficiency claim that this case is meant to challenge.

Daniel's responsibility explains why although he is not blameworthy, some *normative residue* may remain in Daniel's case. By *normative residue* I mean the moral significance that remains because the action is attributable to his agency even though praise or blame is absent. Here, the harm is connected to Daniel's agency in a way that I think may leave him with special reasons to acknowledge the harm, assist the child, or bear some of the resulting costs associated with the child's harm. A further mark of Daniel's responsibility is that it can be understood in terms of *liability to a defensive or preventive burden*. Following McMahan's discussion of liability to defensive harm,² we might say that if the only way to prevent the harm to the child were to impose a necessary and proportionate burden on Daniel, it would be appropriate to do so in a way that would not be appropriate for a mere bystander. This is not because Daniel is blameworthy, but because he stands in a responsibility-relevant relation to the threat. His agency is connected to the risk in a way that the bystander's agency is not.

² Jeff McMahan, "The Basis of Moral Liability to Defensive Killing," *Philosophical Issues* 15 (2005): 386–405, <https://doi.org/10.1111/j.1533-6077.2005.00073.x>.

Someone might still object that the *normative residue* in Daniel's case arises merely from his involvement in a harmful event, rather than from responsibility in the relevant sense. In order to address this matter, I will contrast Daniel's case with one in which the harmful outcome is not attributable to the agent's agency in the same way. Suppose David begins in circumstances similar to Daniel's, but his car is then hijacked by a malicious external agent who takes control of its movement, leaving David unable to do anything about it. If the car then strikes a child, David may still experience shock or sorrow, but he would not bear the same kind of *normative residue* as Daniel does. The reason is that David's connection to the harm is merely causal in a way Daniel's is not. Daniel's residual reasons arise not from involvement in harm as such, but from the fact that the harm is linked to his agency through a knowingly risk-imposing action. This contrast, therefore, supports my claim that the relevant difference lies in moral responsibility, not in causal involvement.

Even though Daniel is morally responsible, he is not morally blameworthy for X. I argue for the claim that appraisability depends not only on an agent's evaluative attitude but also on how those attitudes are expressed and enacted in action and practical reasoning. Praise or blame is fitting only when the agent's conduct, in light of those evaluative attitudes, reveals something that makes the agent praiseworthy or blameworthy. One is praiseworthy when one's action not only reveals moral attentiveness, conscientiousness, integrity, good judgment and serious effort but also manifests them in action. One becomes blameworthy when one's action expresses and manifests things like indifference to moral reasons, negligence, selfishness, poor judgment, and failure to respond appropriately to what she could understand and do. In Daniel's case, his conduct reveals nothing that makes blame fitting. He does not drive recklessly, or negligently. Instead, he drives carefully and follows the rules. The harm, therefore, does not reveal any defect in Daniel's evaluative attitudes or in the manner in which he exercises his capacities.

The bad outcome, in this case, does not express a flawed moral standpoint. Rather, it is a matter of moral luck.

At this point, one might raise an objection based on the issue of moral luck. Issues of moral luck arise when factors outside an agent's control affect the outcome of her action and seem to affect how we morally assess her action.³ In Daniel's case one might say that Daniel is not responsible for *X* because the harmful event depends on factors outside his control, such as the child's sudden appearance. Thomas Nagel emphasizes this point through his control condition, which says that responsibility requires that we are only judged for what is under our control.⁴ But outcomes are often influenced by factors beyond an agent's control. Two equally careful drivers may face different outcomes purely by luck. If responsibility requires control over outcomes, Daniel might seem not to be responsible for the injury because the child's sudden appearance was unforeseeable.⁵ This objection thus highlights a common intuition: *if outcomes matter, luck could undermine responsibility in Daniel's case.*

However, this objection conflates responsibility with outcome alone. Although luck affects whether Daniel's risk-imposing act of driving results in harm, it does not erase his agential connection to *X*. Daniel is responsible for *X* because causing harm to the child through the risk-imposing driving is attributable to him as an exercise of his agency. He knowingly undertook the activity, appreciated the foreseeable risk, and acted under his own control. Nagel also acknowledges that outcomes can be shaped by luck, even though agents can still be responsible for their actions when actions result from their intentions and reasoning.⁶ However,

³ Thomas Nagel, "Moral Luck," *Proceedings of the Aristotelian Society, Supplementary Volumes* 50 (1976): 139–41.

⁴ Nagel, "Moral Luck," 138–41.

⁵ Nagel, "Moral Luck," 142–44.

⁶ Nagel, "Moral Luck," 142–44.

I argue that Daniel remains responsible not because he had a bad intention but because he exercised his agency. He is not blameworthy because the injury does not result from a manifestation of poor evaluative attitudes. Therefore, my answer to the moral luck objection is that the role of luck does not undermine Daniel's responsibility for *X*; it undermines only the inference from harm to moral fault. Responsibility remains fixed at the level of agential attributability, while appraisability depends on whether the agent's evaluative attitudes and their manifestation in conduct make praise or blame fitting. Daniel is not blameworthy, not because *X* is disconnected from him, but because *X* reveals no defect in his evaluative attitudes in the manner in which he exercised his capacities. Daniel's case thus shows that an agent may be responsible for *X* without being appraisable for *X*.

Another challenge arises when the agent's action itself is shaped by unjust social structures. The intended case illustrates how an agent's practical options can be constrained by forces not of her making. If responsibility requires agency, can it survive under structural injustice? Structural injustice occurs when social arrangements or institutional practices create injustice and disadvantage for agents but no one individual is identified as responsible for creating the unjust arrangements.⁷ A case can be presented in the following manner to analyze this matter:

3.3.2 Sara: Structural Injustice

Consider Sara, a government official and the sole provider for her family, who works within a rigid bureaucratic system that administers long-term unjust policies. Sara knows that the policy she helps implement will likely unjustly harm others over time. She also knows that her own ability to resist or alter the system is extremely limited, and that resigning would leave her family without support. Even though she does not support this unjust policy, she continues to perform her role.

⁷ Iris Marion Young, *Responsibility for Justice* (New York: Oxford University Press, 2011), 45, 47, 52.

Unlike Daniel's case, the difficulty here does not arise from moral luck but from structurally constrained agency. The relevant action in Sara's case is Sara's participation in implementing an unjust policy under structural constraint. However, Sara cannot be dismissed as a mere causal instrument of the system. She understands the likely injustice, recognizes her own role in sustaining it, and continues to participate in the institutional process. Her agency is constrained by the system, but it is not totally diminished. The reason is that she is not subject to a form of constraint that makes her unable to understand the moral reasons against participating in the unjust policy. For that reason, her voluntary participation remains attributable to her, and she is responsible for her participation. This point is illuminated by Iris Marion Young's account of structural injustice. Young, and a whole subsequent literature on structural injustice argue that, while individuals cannot be held responsible in the traditional sense for structural injustice, they nevertheless participate in agential ways that generate obligations.⁸ Even though I do not agree with their specific way of understanding agents' responsibility and appraisability under structural injustice, I agree with their key insight that there is a morally significant agential dimension even without the relationship required for appraisability. This thought is important here because it supports the claim that Sara's involvement remains morally significant without forcing the conclusion that she is blameworthy.

I argue that whether Sara is appraisable for her participation depends on the relation between what she takes to matter morally and how that concern is carried through her conduct. Even though Sara believes that the system is unjust, she cannot be praiseworthy just for having

⁸ Young, *Responsibility for Justice*, 95–96, 104–5; Robin Zheng, “What is My Role in Changing the System? A New Model of Responsibility for Structural Injustice,” *Ethical Theory and Moral Practice* 21 (2018): 869–75, <https://doi.org/10.1007/s10677-018-9892-8>.

that thought. The reason is that she is still knowingly involved in carrying out an unjust practice, which disqualifies her from being praiseworthy. Even though she does not endorse the unjust policy or act out of indifference to those who will be harmed by it, she continues to participate. She is also not blameworthy because the constraints under which she acts affect what her participation manifests about her evaluative attitudes. A different case would arise if Sara used her position to take significant personal risks in order to resist, expose, or substantially reduce the injustice. In such a case, her conduct might express a degree of courage or integrity sufficient for praise. But the present case is not of that sort. Her conduct in the original case, therefore, occupies a morally burdened middle space: she remains responsible for what she does, but she is neither praiseworthy nor blameworthy. This does not mean that Sara's conduct cannot be morally assessed in any broader sense. One might say that her conduct receives a neutral assessment once the constraints on her agency are taken into account. But in the sense of moral appraisability used in this thesis, to be appraisable is to be praiseworthy or blameworthy. Since Sara's participation does not make either praise or blame fitting, she is responsible without being morally appraisable in that sense.

Alternatively, we can imagine another person who works in the same bureaucratic system but, unlike Sara, endorses that unjust policy and participates out of commitment to its aims. In such a case, the person is not only responsible but also blameworthy for her participation. The reason is that her conduct expresses and enacts a morally defective evaluative attitude in the very way she deliberates, chooses, and acts. Sara's case is different because she neither endorses nor supports the unjust policy. Her participation is constrained and accompanied by an awareness of the reasons against it. Sara's participation takes place under pressure, limited options, and constrained practical circumstances. These facts do not automatically remove blame. Rather, they affect what her participation manifests about her evaluative attitudes. If her participation manifested endorsement, she would be blameworthy.

But in the case described, Sara's constrained participation does not manifest that kind of evaluative attitude. For that reason, she is responsible for what she does because she understands the injustice, recognizes her role in the policy, and continues to participate voluntarily within the limits of her situation, yet she is not blameworthy for it. Sara's case thus illustrates a different route by which responsibility and appraisability can come apart.

3.4 What the Cases Show

The cases of Daniel and Sara show that an agent can be responsible for *X* without being praiseworthy or blameworthy for *X* in different ways. In Daniel's case, the divergence arises from moral luck: the harmful event is linked to Daniel's agency; nevertheless, the outcome does not reveal any moral fault in his conduct. In Sara's case, by contrast, the divergence arises from structural constraint: her participation in an unjust practice remains attributable to her, yet the constraints under which she acts make praise and blame inappropriate. Therefore, together they disprove the claim that responsibility is sufficient for appraisability.

These cases also suggest a more general pattern for generating counterexamples to the biconditional that links responsibility and appraisability. Such cases may arise when the agent's conduct remains attributable to her agency in a morally significant way so that responsibility is preserved and the agent's normative residues can be preserved, but some further feature prevents appraisal from following. In Daniel's case, that further feature is moral luck, and in Sara's case, it is structural constraint. Therefore, the result is a class of cases in which agents remain responsible without thereby becoming appropriate objects of praise or blame.

At this point, an intuitive worry may arise regarding the importance of moral responsibility in our everyday lives. We may worry that if responsibility is separated from appraisal, it may lose its significance. Moreover, if an agent is neither praiseworthy nor blameworthy, what would remain of responsibility at all? My claim is that what remains is not

trivial. Responsibility still marks an important moral relation between agents and actions because it identifies the action as attributable to the agent in a way that goes beyond mere causal involvement. This matters in everyday life because responsibility can still ground demands of acknowledgment, repair, compensation, or cost-bearing, even when praise or blame is not fitting. As I have discussed in Daniel's case, responsibility may leave behind a kind of *normative residue*: he may owe compensation, bear certain costs, or experience a special form of regret connected to the harm. Sara's case shows something similar in a different register: although blame is unstable, her participation remains morally significant because it is still connected to her agency and not merely to the surrounding structure. And I think that her responsibility may leave her with special reasons to acknowledge her role in the unjust process, or to answer for her participation. Responsibility, in this sense, retains its significance even where praise or blame is absent, unstable, or misplaced.

3.5 Conclusion

In this chapter, I have shown that responsibility is not sufficient for appraisability, understood as praiseworthiness or blameworthiness. I have presented two moral cases that show that an agent can be responsible for *X* without being praiseworthy or blameworthy for *X*. The next chapter will shift focus from these counterexamples to a constructive analysis. I will provide the *conditions of responsibility and appraisability* in detail. My aim will be to show precisely what it takes for an agent to be responsible and what it takes for her to be praiseworthy or blameworthy. This analysis will aid us in completing the defense of the *asymmetry thesis*: *responsibility is necessary but not sufficient for appraisability*. Having articulated the conditions of both responsibility and appraisability, I will analyze whether responsibility is necessary for appraisability, and I will argue that it is.

—

Chapter Four

A Constructive Account of Moral Responsibility and Moral Appraisability

4.1 Introduction

In the previous chapter, I made an intuitive case, through Daniel and Sara, for the claim that moral responsibility is not sufficient for moral appraisability. But that intuitive case, while it may be useful in supporting the claim, does not yet explain how moral responsibility may be present without moral appraisability, understood as praiseworthiness or blameworthiness. To explain this, we need a constructive account of the conditions under which agents are morally responsible and morally appraisable.

My aim in this chapter is to provide the constructive account that the previous chapter left open. I will proceed in three steps. *First*, I will explain moral responsibility as concerning the agential connection between an agent and an action: whether the action is attributable to the agent as an exercise of her agency. This connection obtains through the agent's required control and relevant awareness. *Second*, I will explain moral appraisability as concerning whether the agent is praiseworthy or blameworthy in virtue of how the agent's evaluative attitudes are expressed and enacted in action. The required conditions for moral appraisability are the agent's evaluative attitudes and the manifestation of those attitudes in deliberation, choice and action. *Finally*, I will bring these accounts together to show why moral responsibility is necessary for moral appraisability, even though it is not sufficient for it. In this way, the chapter completes the asymmetry I have defended in this thesis.

4.2 Moral Responsibility: a Distinct Form of Agential Connection

An agent may be related to an action in at least three different ways. She may be causally linked to it, morally responsible for it, or morally appraisable for it. These three should not be collapsed into one another since they are concerned with different questions. Causal connection asks whether an agent has some role in bringing about an action. Appraisability asks whether praise or blame is fitting with respect to a particular action; that is, whether the agent is praiseworthy or blameworthy. Moral responsibility, as I understand it, stands between these two. It asks whether an action is attributable to an agent as an exercise of her agency so that the action counts as her own in the morally relevant sense, that is, in the sense required for moral responsibility.

This distinction matters because moral responsibility cannot be reduced either to causal connection or to moral appraisability. Causal connection is not sufficient because an agent may be causally connected to an action through an accident, manipulation, or blameless ignorance without the action being hers in the morally relevant sense. On the other hand, moral appraisability requires something further than moral responsibility. An agent may be morally responsible for performing a routine action, such as carrying out an ordinary work task, without being praiseworthy or blameworthy for it. This does not mean that an agent cannot be morally assessed at all. As I have said before, an agent can be morally assessed in a broad sense and receive a neutral or lukewarm judgment. However, the sense of moral appraisability that I have used in this thesis requires more than neutral assessment. In this sense, to be morally appraisable, the action for which the agent becomes praiseworthy or blameworthy must reveal something morally significant about the agent's exercise of agency. So, for a routine action, an agent may be open to moral assessment in a broad sense but she would not necessarily be praiseworthy or blameworthy. The cases discussed in the previous chapter can be mentioned here to clarify this distinction. Daniel, for example, is

not merely causally connected to the harm, since he knowingly drives in hazardous weather and exercises control over what he does. The harm, therefore, is connected to him through the exercise of his agency. However, this does not make Daniel blameworthy because his conduct does not reveal a defective evaluative attitude that would make blame fitting. Sara's case also highlights the same point in a different way. Whereas Daniel's case involves risk and luck, Sara's involves agency exercised under constraint. Her constrained participation in an unjust bureaucratic system may still be connected to her in the sense required for moral responsibility. Although it is something she does within her constrained agency, the constraints under which she acts limit how her evaluative attitudes are expressed in action, thereby making blame inappropriate. These cases show that there is a form of *agential connection* between agents and actions that is stronger than mere causal involvement but not yet sufficient for praiseworthiness or blameworthiness. This connection is what moral responsibility marks.

To explain the connection, we can now briefly recall my concept of moral responsibility: an agent is morally responsible for action *X* when *X* is attributable to her as an exercise of her agency. Here, in my concept, *attributability* names the agent–action relation that moral responsibility requires: the relation in which an action counts as the agent's own in the morally relevant sense. I argue that **control** and **relevant awareness** are what must be in place for attributability to obtain. They are individually necessary because without control, the action cannot proceed through the agent's guidance, and without relevant awareness, the agent is not a knowing participant in what she is doing under its morally relevant description. Together these two are sufficient because when an action proceeds through the agent's guidance and the agent has the relevant awareness of what she is doing, the action counts as attributable to the agent as an exercise of her agency. Thus, control and relevant awareness are not separate conditions added alongside

attributability; rather, they explain when the responsibility-relevant relation between agent and action is present.

4.2.1 Attributability: the Agent–Action Relation Required for Moral Responsibility

In my account of moral responsibility, *attributability* explains why an action can be regarded as the agent’s deed, rather than merely as something connected to her causal history. Now, I need to clarify what I mean by *attributability*.¹ To say that an action *X* is attributable to an agent *S* is to say that *X* counts as *S*’s own action in the morally relevant sense. Again, this does not mean merely that *X* is caused by *S* or that *X* occurred through *S*’s bodily movement. Instead, for *X* to be attributable to *S*, it must issue from *S* as an exercise of agency. That means that attributability includes both doing and owning. The action must be something the agent does, but it must also count as hers in the morally relevant sense.² By “owning” an action, I do not mean owning it as a possession. I mean that the action stands in the right relation to the agent’s agency. It springs from her as an exercise of agency instead of merely occurring through her. That means that the action must proceed through the agent’s own control and relevant awareness. Only then can the action be attributable to her in the sense required for moral responsibility.

¹ My use of attributability is narrower than some familiar uses of the term in the responsibility literature. In accounts associated with Watson, Smith, Scanlon, and Wolf, attributability is often used to ask whether an action or attitude expresses something about the agent’s evaluative standpoint. My concern here is different. I use attributability to mark the agent–action relation required for moral responsibility: whether the action counts as the agent’s own exercise of agency. Since my concern is this responsibility-relevant relation, attributability on my account cannot be separated from control and awareness.

² Here, “doing” includes both actions and omissions. An omission can be attributable to an agent when the failure to act issues from her agency in the relevant way.

A simple case can be used to demonstrate why the substantial agential connection between agents and actions is needed for moral responsibility. Suppose an agent's action outwardly expresses hostility, but the action results from covert manipulation or from a severely disordered mental state. The agent lacks control and relevant awareness. What is missing, in this example, is the right agential relation. What remains, then, is to explain how attributability is secured. In the next two sections, I will argue that it is obtained through the agent's **required control** and **relevant awareness**.

4.2.2 Required Control

If attributability is the required relation between an agent and an action, we need to know what needs to be satisfied for that relation to obtain. I think that it requires control. An action cannot count as the agent's own if it is not done under the agent's guidance. It may happen that an agent merely stands in the causal history of an action, but she lacks control over the action since her body is forcefully moved, mechanically directed, or otherwise used in an instrumental way that bypasses her agency. What these instances show is not simply that the agent lacks the capacity required to do the action but also that the action does not proceed under the agent's guidance. If control is not required for attributability, then we have to consider agents responsible in cases of the above types. Again, this control condition, as I understand it, is not primarily concerned with whether the agent could have done otherwise. The reason is that the issue relevant for attributability is not whether the agent had open alternatives, but whether the action performed proceeds through her practical and deliberative agency. Suppose an agent decides to go shopping. After considering her reasons, she chooses one route to the mall. Unbeknownst to her, the other routes to that mall are closed that day, so the route she takes is in fact the only available route. In this sense, she could not have taken a different route. Still, the route she actually takes proceeds through her own

deliberation and guidance. This shows why the relevant form of control concerns how the action actually issues from the agent, not whether alternative possibilities were available.

Control, in the required sense, must first include the physical capacity to perform the action. By *physical capacity*, I mean that an agent must be able to initiate and carry out the action in the world, rather than merely desiring or intending it. Suppose an agent intends to sign a document, but her hands are temporarily paralyzed. Her intention is not enough for her to control the action of signing the document because she lacks the physical capacity to carry it out. This shows why desire or intention alone is not enough for control; an agent also needs relevant physical capacity to perform the action.

However, physical capacity alone is not sufficient. An agent may be physically capable of performing an action and still fail to control it in the sense responsibility requires. This might happen when the agent cannot guide what she is doing through practical thought. For this reason, control must also include deliberative capacity. By *deliberative capacity*, I mean the capacity to direct one's action through practical reasoning that meets at least a basic level of rationality. By a basic level of rationality, I understand that an agent can understand relevant considerations and guide her action accordingly. It includes the capacities to consider what one is doing, respond to relevant reasons, and guide one's action accordingly. To consider what one is doing means to have some grasp of the action one is performing. To respond to relevant reasons is to be able to recognize considerations that count for or against acting in that way. To guide one's action is to allow those considerations to shape what one does. Without this capacity, the agent may still perform the relevant bodily movements, but the action would not proceed from her as a guided exercise of agency. While physical capacity explains how the agent is able to carry out the action, deliberative capacity explains how the action is guided by her agency rather than merely occurring

through her. Both are needed for the action to count as controlled in the sense required for attributability.

At the same time, the control required for attributability should not mean that an agent with control always acts under ideal or unconstrained conditions. If having control means acting under ideal and unconstrained conditions, very few actions would count as actions over which agents have control. For agents often act with moral responsibility despite difficulty, pressure, limited options, or unjust circumstances. This is also why my account is not primarily concerned with the ability to do otherwise. Therefore, the relevant concern is not whether the agent acts in perfectly unconstrained conditions, but whether the action still proceeds through her own practical and deliberative capacity and remains an action that the agent is able to perform. This is why it is important to distinguish between cases in which agency is bypassed and cases in which agency is pressured or constrained. We know that being bodily overpowered, mechanically directed, or otherwise made into an instrument through which an event occurs destroys the agent's control altogether. This is because in those cases the action does not proceed through the agent's own guidance even though her physical capacity has been used. Other forms of interference, however, do not eliminate control in the same way. They may narrow down an agent's options, impose serious pressure or make the circumstances unjust, but the agent may still act through her deliberative capacities.

Coercion by threat is an important example of this second kind of case. Carla Bagnoli argues that coercion by threat does not simply destroy agency, because the threatened person is still addressed as a deliberating agent. The threat also works by giving the agent reasons, even

though those reasons are imposed in a deeply unjust way.³ The coerced agent therefore still acts through her deliberative powers rather than being moved in the way a body is moved by *brute force*.⁴ Robert Campbell makes a related point when he argues that duress does not remove the agent's control altogether, even though the surrounding circumstances and options are not of her own making.⁵ The agent remains responsible for the choice she makes, even if she is not responsible for the circumstances or options imposed upon her.⁶ What matters for my account is that such cases do not involve the same kind of agency-bypassing found in brute force or bodily manipulation.

If my account is plausible, this is also where the distinction between responsibility and appraisability becomes especially helpful. If we did not distinguish between responsibility and appraisability, we would have to say either that coerced agents were appraisable or that they were only causally responsible. But this mischaracterizes their situation. My account allows us to say that an agent may remain responsible for a choice made under pressure because the choice still proceeds through her own guided agency, without thereby having to claim that she is praiseworthy or blameworthy for making it. Responsibility remains because the action is still attributable to her; appraisal may nevertheless be inappropriate because of the conditions under which the action is

³ Carla Bagnoli, "Claiming Responsibility for Action Under Duress," *Ethical Theory and Moral Practice* 21, no. 4 (2018): 857, <https://doi.org/10.1007/s10677-018-9904-8>.

⁴ This is the kind of force the agent cannot control. When one is harmed by lightning or by any natural phenomenon that event is an example of brute force.

⁵ I am not using "duress" to mark a separate category here; rather, I take it to refer to the same general kind of pressured agency, where the agent's options are shaped by threat or severe constraint without her agency being bypassed altogether.

⁶ Robert Campbell, "Duress and Responsibility for Action," *Journal of Applied Philosophy* 1, no. 1 (1984): 138–39.

performed. This point also guides us to explain why Sara, the agent from the previous chapter, is responsible but is not blameworthy. We know that her options are sharply constrained by unjust structural conditions, but her agency is not bypassed. She still makes a choice through her practical and deliberative capacities, even when she does so under pressure. Her action may therefore remain attributable to her, even where praise or blame is complicated or misplaced. Control, then, should exclude bypassed action, but not every constrained choice.

The control condition can therefore be stated as follows: if action *X* is to be attributable to agent *S*, then *S* must possess the physical capacity to perform *X* and the deliberative capacity to guide *X* through minimally rational practical thought. This condition explains one part of what is needed for an action to count as the agent's own in the sense required for moral responsibility. However, an agent may possess the physical and deliberative capacities required to guide an action and yet fail to be responsible for it if she lacks relevant awareness of what she is doing or of the morally relevant features of the situation. This is why attributability requires not only control but also awareness.

4.2.3 Relevant Awareness

Awareness is the second grounding condition of attributability. The importance of awareness as a condition of responsibility has been discussed in the literature on culpable and blameless ignorance.⁷ Including awareness as a condition of responsibility is not unique to my account of responsibility. But the distinctive feature of my account is the way I have used it in explaining

⁷ For discussions of ignorance, awareness, and moral responsibility see Holly Smith, "Culpable Ignorance," *The Philosophical Review* 92, no. 4 (1983): 543–71; Gideon Rosen, "Culpability and Ignorance," *Proceedings of the Aristotelian Society* 103, no. 1 (2003): 61–84; Michael Zimmerman, "Moral Responsibility and Ignorance," *Ethics* 107, no. 3 (1997): 410–26.

when an action is attributable to an agent in the sense required for moral responsibility. On my view, if an action *X* is to be attributable to agent *S*, then *S* must have relevant awareness of what she is doing and of the morally relevant features of the situation in which she does *X*. I use “relevant” to indicate that awareness does not mean having complete moral or factual knowledge of every fact connected with the action. Moreover, it does not demand that the agent anticipate every possible consequence. For the purpose of my thesis, the awareness condition is modest because it only requires that agents understand what they are doing and have required awareness of the morally relevant features of the situation in which they act for the action to count as theirs. It is not required that agents must know every probable consequence of their actions.

An example will help us see the point more clearly. Suppose Fiona presses a button every morning believing rationally that it grinds coffee. Unbeknownst to her, one morning the button is rewired to burn a house next door. Since she is pressing the button with no one constraining her, she meets the control condition. She can reason about whether or not to press the button, and, given her interest in making coffee and her evidence about what the button does, pressing the button makes sense from her point of view. But she fails the awareness condition because she does not have the right information. That is, she does not know that by pressing the button she is going to burn down a house. She is causally linked to the action, but burning the house is not attributable to her as her deed in the sense required for moral responsibility. This case shows why relevant awareness is required in addition to required control. Although control explains how the action passes through the agent’s physical and deliberative capacities, it cannot by itself explain whether the agent is a knowing participant in what she does. Fiona’s case, I think, is understood as a case of factual blameless ignorance, which also highlights in which situations an action can be considered a result of blameless ignorance that undermines responsibility. Later in this chapter, I

will consider how the awareness condition should be understood when the agent's ignorance is moral rather than merely factual. The purpose is to highlight that moral uncertainty should not be treated in exactly the same way as ordinary factual ignorance such as Fiona's ignorance in this case. For now, Fiona's case helps show the basic point that a blameless lack of factual awareness can interrupt the responsibility-required connection between agents and actions.

Blameless ignorance can interrupt responsibility because it prevents the action from being attributable to the agent in the required sense. If the agent could not reasonably have known what she was doing or could not reasonably understand the morally relevant features of the situation, then the action cannot be considered issuing from her as a knowing exercise of her agency. The reason is that ignorance undermines her ability to guide her action through her deliberative capacities. In such cases, the lack of awareness breaks the agential connection required for moral responsibility. Fiona's lack of awareness is not a result of negligence, avoidance, or failure to inquire. She could not reasonably have known that the button had been rewired. Also, there is no prior failure of attention, inquiry, or understanding on Fiona's part that could make the burning of the house attributable to her in the sense required for moral responsibility. Fiona's action therefore falls under the category of blameless ignorance. This case differs from cases of negligence, avoidance or failure to inquire where the attributability relation is maintained. The reason is that in those cases the ignorance remains connected to an agent's earlier failure to attend to or inquire into the relevant facts when she had reason and opportunity to do so.

Carolina Sartorio's discussion concerning this matter can be mentioned here. She says that responsibility depends partly on whether the relevant character of what the agent does falls within

the scope of what she could reasonably understand or foresee.⁸ If the morally relevant character of the action lies wholly outside the agent's awareness, and could not reasonably have been within it, then the action does not reflect her agency in the way responsibility requires. On my view, the reason is that the agent cannot guide her action according to the feature that she could not reasonably recognize. The action may still be connected to her but it does not issue from her as a knowing exercise of her agency. This, again, supports the claim that awareness is not an optional addition to responsibility, but one of the grounding conditions under which an action becomes attributable to the agent.

However, not every lack of actual awareness undermines responsibility in the same way explained above. The awareness condition should not be understood as requiring actual awareness in every case without qualification. In the case of awareness, the relevant concern is to distinguish between *blameless ignorance* and *culpable ignorance*. The distinction can be explained by using the reasonable expectation standard. According to the reasonable expectation standard, if an agent could not reasonably have been expected to know or understand the morally relevant features of the situation in which she acts, then her ignorance is blameless. But if she could reasonably have been expected to know or inquire into those features, then her ignorance is attributable to her.⁹ Therefore, we can say that blameless ignorance may interrupt agents' responsibility by interrupting attributability, whereas culpable ignorance preserves responsibility because the lack of awareness

⁸ Carolina Sartorio, "Causation and Responsibility," *Philosophy Compass* 2, no. 5 (2007): 751.

⁹ I am borrowing Sartorio's reasonable expectation idea, but I am using it for my own purpose which is to explain when ignorance does or does not block attributability.

is traceable to the agent's earlier failure to attend, inquire, or understand when she had the capacity, fair opportunity, and sufficient reason to do so.

The relevant point is not that the agent should have known better, but that her failure to know can itself be part of what connects the later action to her agency. Since the ignorance is traceable to the agent's earlier failure, it does not stand as an external constraint that blocks attributing the later action to the agent. Rather, it becomes part of the route through which the later action remains connected to her agency. The clearest cases of this kind happen when the agent had the capacity and fair opportunity to understand the relevant features of the situation, had sufficient reason to attend to them, but she failed to do so. Such failure may show carelessness, avoidance, and disregard when the situation reasonably required her to do otherwise. This does not mean that mere ignorance is irrelevant to moral responsibility. But the point is that, in cases of culpable ignorance, the ignorance is connected to the agent's earlier exercise of or failure to exercise her agency. The earlier failure can become part of the agential history of the later action. In such cases, the later ignorance resulting from the earlier failure does not stand outside the agent's agency in the way blameless ignorance does. Thus, the responsibility-relevant connection between agents and actions is preserved, and the agent becomes morally responsible.

However, one might worry about how far back in the temporal chain of an agent's actions or decisions we need to look when assessing a particular action. Even though I cannot give a complete account of how far such tracing should go, I think that not every earlier action or decision needs to be included for the assessment of the action. I think that only earlier actions or decisions that help explain the agent's later lack of awareness in the responsibility relevant sense are relevant for such tracing. Some actions or decisions of an agent might be irrelevant to the particular action

we are assessing. For an action or decision to be relevant, the agent in question must have had the relevant capacity, fair opportunity, and sufficient reason to attend, inquire, or understand.

We can consider an example. Suppose an agent fails to read a warning label. If the agent fails to read the warning label even though she had reason to read it, or deliberately avoids finding out facts that would reveal the moral significance of what she is doing, her later ignorance does not stand outside her agency. If she later gives someone the wrong dose because of her earlier ignorance, then the ignorance would be attributable to her. This is because it is produced or sustained by her own failure to attend, inquire, or take seriously facts that she had reason to understand. This, again, does not imply that the general ability to know or understand something is enough for responsibility. Nor is it enough that the agent could have chosen to inquire. An agent may have been able to learn some fact and may permissibly have chosen not to do so. In that kind of case, the ignorance may be connected to her agency in a broad sense, but it is not attributable to her in the responsibility-relevant sense. Many things are knowable in principle, but not all of them are things an agent could reasonably be expected to know in the circumstances. What matters is not bare possibility, but whether the agent had the relevant capacity, fair opportunity and sufficient reason to understand. By *relevant capacity*, *fair opportunity* and *sufficient reason*, I mean that the agent was capable of grasping the relevant facts or moral features, the circumstances gave her a reasonable chance to notice, inquire, or understand, and there was something in the situation that should have prompted her to attend more carefully. Only when these conditions are present does the failure to understand remain connected to the agent's agency in the right way.

With this, one may ask: why is the agent expected to exercise this kind of attention or inquiry? It is not that the agent already has the relevant information. If she had it, she would not be ignorant in the relevant sense. Rather, the point is that the circumstances may give her sufficient

reason and fair opportunity to attend, inquire, or understand. She could read the warning with a reasonable understanding. She could also understand the other features that should prompt her attention. Additionally, she was not severely constrained or manipulated in a way that she could not inquire about the relevant facts before or while giving the dose. When those conditions are present, her failure to know is not merely a lack of information; it is traceable to the failure to exercise her agency. If the relevant fact could not reasonably have been foreseen or understood, the later action is not attributable to her under that morally relevant description. But if she had reason and opportunity to understand and failed to do so, the ignorance becomes part of the agential history of the later action.

The above point also applies to control. A blameless lack of control interrupts responsibility, as in cases where an agent's body is forcibly moved or her agency is bypassed. But a lack of control that is traceable to an earlier controlled choice or failure may not interrupt responsibility in the same way. In such cases, responsibility is not grounded in the later lack of control by itself but in the earlier exercise or failure of agency that explains why the later lack of control occurred. Control and relevant awareness are therefore structurally similar. In both cases, a blameless absence may interrupt responsibility, while an absence traceable to the agent's earlier agency may preserve the responsibility-relevant connection.

The examples discussed so far are cases of factual or descriptive ignorance in which the agent does not know what she is doing, what her action will cause, or what facts are relevant to the situation. However, there are other types of ignorance in which an agent knows what she is doing but is uncertain or even reasonably mistaken about what morality requires. The awareness condition should not be understood as requiring a fully correct moral theory or complete certainty about what is morally required, permitted, or supererogatory. For example, a parent may

reasonably but mistakenly believe that a certain way of shaping a child's worldview is permissible. If the parent understands what she is doing, recognizes the morally significant features of the situation, and guides her action through deliberation, then the action is attributable to her even though what she thinks is permissible is mistaken. The moral error does not automatically remove responsibility. It may affect the appropriateness of praise or blame. But this again does not mean that an agent's responsibility is never undermined in cases of moral ignorance. This happens when the agent's moral ignorance is so deep and non-culpable that she cannot reasonably understand the morally relevant nature of what she is doing because she lacks a fair opportunity to understand.¹⁰

With this clarification in place, the awareness condition can finally be stated as follows: if action *X* is to be attributable to agent *S*, then either *S* must possess the relevant awareness of what she is doing and of the morally relevant features of the situation in which she does *X*, or her lack of such awareness must itself be attributable to her through an earlier failure to attend, inquire, or understand when she had the capacity, fair opportunity, and sufficient reason to do so. Relevant awareness does not require complete factual knowledge or a fully correct moral theory. The awareness condition is satisfied either when the agent has the relevant awareness at the time of action or when her lack of awareness is attributable to an earlier failure of her agency. Together

¹⁰ Susan Wolf's JoJo case can be read as an example of this type, where JoJo, the son of an oppressor, never got the chance to know what morality requires. To briefly state: JoJo was so influenced by his dictator father that he never had the chance to realize that torturing people could be a wrong thing to do. He followed his father afterwards. In this case, JoJo's action is not attributable to his agency as he lacks required understanding of what he does. Therefore, JoJo is not responsible for his moral ignorance. See Susan Wolf, *Freedom Within Reason* (Oxford: Oxford University Press, 1990), 45.

with the control condition, this explains how an action becomes the agent's own in the sense required for moral responsibility.

4.2.4 The Final Structure of Moral Responsibility

The final version of my account of moral responsibility can now be stated more precisely. On my view, what it is for an agent *S* to be morally responsible for an action *X* is for *X* to be attributable to *S* as an exercise of *S*'s agency. Attributability is the responsibility-relevant relation between agents and actions: the relation in which the action counts as the agent's own in the morally relevant sense. Control and the awareness conditions are individually necessary and jointly sufficient conditions of attributability. The awareness condition may be fulfilled either by the agent's relevant awareness at the time of acting or by an earlier attributable failure that explains the agent's lack of awareness. That is, action *X* is attributable to agent *S* if and only if *S* has the required control over *X* and either *S* possesses the relevant awareness or lacks that awareness because of an earlier attributable failure to exercise her agency.

This account of attributability matters because we care about whether an action is really the agent's. This matters in its own right, not just because praise or blame may later follow. When an action is attributable to an agent, it stands in a substantial relation to the agent's agency. Even when the issue of praise or blame does not arise, an action may still stand in a morally important relation to the agent in virtue of the agent's moral responsibility for the action. That relation may ground acknowledgment, regret, repair, or reasons to bear certain burdens connected with what one has done. Daniel's case can be recalled here. Although Daniel is not blameworthy, the harm remains connected to his agency in a way that may ground acknowledgment of the harm and reasons to assist or bear some of the costs. This is different from the button-pressing case discussed in the awareness condition. In that case, Fiona is causally connected to the burning event but that

event is not attributable to her because, as we have seen, she is blamelessly ignorant of what pressing the button will do. When attributability is absent, moral responsibility related responses such as acknowledgement, repair, or compensation are not required on the responsibility relevant ground. Fiona may feel sympathy or she may even regret what happened as a result of her causal involvement. Her causal involvement in the relevant event gives her a special reason to consider herself differently from a completely uninvolved bystander. But that connection, at the same time, is not sufficient for praise or blame. If Fiona thinks that she owes compensation because the burning event is morally attributable to her, then she is mistaken. This is a significant strength of my account of moral responsibility. Without distinguishing moral responsibility from both causal connection and appraisability, it would be difficult to explain why such responses may be fitting. If Daniel were merely causally connected to the harm, the normative residues would seem ungrounded; but if responsibility were equivalent to appraisability, then recognizing his responsibility would seem to imply blame. My account avoids both results. Responsibility still marks a substantive moral relation between agents and actions, even where appraisal does not follow.

4.3 Moral Appraisability: Appraising the Agent for an Action

The above discussion of responsibility emphasizes what conditions are needed to be satisfied for an agent to count as a morally responsible agent for a particular action. However, examples from the previous chapter showed that moral responsibility is not sufficient for moral appraisability. We need further conditions under which an agent can be morally praiseworthy or blameworthy. To recall briefly, on my view, moral appraisability is an assessment of whether an agent is praiseworthy or blameworthy for a particular action. The agent is the primary target of praise and blame, while the action is the basis on which the agent is appraised. We may describe an action as

praiseworthy or blameworthy, but only derivatively. When we say an action is praiseworthy or blameworthy, we mean that by performing the action, the agent manifests something morally significant about her exercise of agency for which she is being appraised. Suppose someone saves a trapped child from a burning house. The agent is praiseworthy for saving the child because the action, saving the child, expresses her courage, care, or moral attentiveness. Therefore, it seems plausible to say that the relation between assessing an action and appraising an agent is that the action provides the basis for appraising the agent in light of what the action reveals about her exercise of agency.

If appraisal concerns whether an agent is praiseworthy or blameworthy for doing an action, then the next question is what the action must reveal about the agent for praise or blame to be fitting. I think that appraisal depends on the agent's evaluative attitudes and how those are expressed and enacted in deliberation, choice, and action so that the action reveals something morally significant about her exercise of agency. By *morally significant*, I mean something morally important that bears on whether praise or blame is fitting for the agent in question. Therefore, we can claim that while moral responsibility asks whether an action is attributable to an agent, moral appraisability asks what, if anything, the action reveals that makes praise or blame fitting. In this section, I will turn my focus to the issue of moral appraisability.

4.3.1 Evaluative Attitudes: the First Condition of Moral Appraisability

Suppose Mary's husband, Jonathan, prepares their children for school and takes them there in the morning. He understands what he is doing and does it without pressure from anyone. The action thus seems attributable to him in the sense required for moral responsibility. But now we can imagine two versions of this case. In one, he takes himself to be merely doing Mary a favor. In the other, he takes the action to be something that morally matters to him. The outward conduct is the

same in both scenarios, but the action is appraisable in a different way.¹¹ The difference lies in how he takes the action to matter. If he treats it merely as doing Mary a favor, the action reflects one evaluative standpoint; if he treats it as part of their shared parental responsibility to care for their children, it reflects another. This demonstrates that outward action alone is not sufficient for moral appraisability. For praise or blame to be fitting, the action must involve the agent's evaluative standpoint toward what she is doing and the morally significant features of the situation. The first condition of appraisability, therefore, is that the agent must have evaluative attitudes toward the action and the morally significant features of the situation. This Jonathan case also helps us to separate two connected questions about appraisability. The questions are: how should we appraise agents? And, are they eligible for praise or blame? My main concern in this thesis has been whether an agent is eligible to be praised or blamed for an action. Once appraisal is fitting, evaluative attitudes also explain what particular appraisal the agent merits.

Recall that I use *evaluative attitudes* to refer to the attitudes through which agents take certain morally significant features of the action or situation to matter. These attitudes may include explicit or implicit moral beliefs, such as beliefs about what matters morally in the situation or what the agent has reason to do. Beliefs matter because they help show how agents understand the moral significance of what they are doing. But they are not exhausted by belief. They may also include care, concern, valuing, attention, and responsiveness. What unifies these attitudes is that they reveal what the agent takes to matter in acting.

¹¹ This case is motivated by James Grant's Father case even though I have modified it to serve my purpose. See James Grant, "Moral Worth and Moral Belief," *Ethics* 133, no. 2 (2023): 220, <https://doi.org/10.1086/722120>.

The reason for understanding evaluative attitudes broadly is that agents' evaluative attitudes cannot always be explained by belief alone. A person may believe that something is valuable without caring about it, and a person may care about something in ways that are not reducible to an explicit or implicit judgment. If appraisability were limited to belief, it would miss some of the ways agents reveal what they take to matter morally. An agent may act with care or responsiveness without forming a clear belief that her action is morally right and may nevertheless be praiseworthy for that. Similarly, an agent may be blameworthy because her actions reveal selfishness, negligence, or failure to attend to what matters, even where the issue is not explained in terms of a false moral belief.

Huckleberry Finn is a familiar example in the discussions of consciousness, moral worth, and inverse akrasia.¹² This case shows the usefulness of understanding the broader aspects of evaluative attitudes. Huck is a free white boy in the antebellum American South who has been socialized to regard racial slavery as normal and to think of Jim, an enslaved man, as someone else's property. Because of this socialization, Huck explicitly believes that helping Jim escape is wrong. Yet Huck goes on to help Jim escape despite his judgment about the slavery system. Now, the question is: if appraisability were restricted to moral belief, what should be said about cases such as Huck in which an agent's explicit moral beliefs are mistaken and conflict with what he does, yet the agent still seems morally praiseworthy? If evaluative attitudes are limited to beliefs, then Huck would fall outside the scope of positive appraisal. But I think Huck is praiseworthy for helping Jim escape. Huck's explicit belief is one part of his evaluative attitude; it does not exhaust that standpoint. In helping Jim, Huck responds to morally significant features of the situation, even

¹² Nomay Arpaly, *Unprincipled Virtue: An Inquiry into Moral Agency* (Oxford: Oxford University Press, 2003), 75–76.

if he cannot fully articulate them in correct moral terms. We may still criticize Huck for his belief, but in appraising Huck for helping Jim, the relevant question is what the action reveals about his evaluative standpoint. His action reveals care, responsiveness, and recognition of Jim's humanity, even if these are not captured in his explicit belief. The point, then, is not that belief is irrelevant to appraisal, but that explicit moral belief may fail to capture the evaluative standpoint from which the agent acts. Thus, it seems reasonable to consider the wider aspect of evaluative attitudes to recognize what agents take to matter morally. At the same time, broadening evaluative attitudes needs clarification. Broadening the category does not mean that every attitude an agent has while acting is relevant to appraisability. Only attitudes that bear on how the agent takes the morally significant features of the action or situation to matter are relevant to moral appraisability.

The first condition, then, identifies the evaluative standpoint from which the agent acts. Since moral appraisability concerns the agent for a particular action, we also need to know how that standpoint is carried through in the action. Two agents may share a similar evaluative attitude and yet differ in appraisability because of the way each acts from that attitude. For example, two agents may both take it to matter morally that people who are trapped in a burning building need help. One may respond with serious attention, practical care, and an effort to help in a way that is appropriate to the situation. Another may recognize the same moral importance but act carelessly, ignore available safe ways of helping, or fail to take the situation seriously. The evaluative attitude may be similar, but what the action reveals about the agent's exercise of agency is different. This shows why appraisability requires a second condition: the agent's evaluative attitudes must be manifested and enacted in deliberation, choice, and action.

4.3.2 Manifestation: the Second Condition of Moral Appraisability

The second condition of appraisability is manifestation. By *manifestation*, I mean the way an agent's evaluative attitudes are expressed and enacted through how she understands the situation, what she decides to do, and how she carries out the action. The need for this condition follows from the first condition. In the first condition, we have seen that evaluative attitudes identify the standpoint from which the agent acts, but they are not sufficient for moral appraisability by themselves. Since moral appraisability concerns whether an agent is praiseworthy or blameworthy for an action, those attitudes must be manifested through the agent's deliberation, choice, and action. This is why evaluative attitudes considered in isolation do not settle whether praise or blame is fitting for a particular action. However, I do not deny that beliefs, attitudes, omissions, or character traits may sometimes be appraised in their own right. A person may be criticized for a belief when that belief reflects prejudice, negligence, disregard, or a failure to attend to relevant reasons. However, my focus for this project is narrower: when the issue is appraisability for an action, the relevant attitudes must be connected to that action through deliberation, choice, and conduct.

Daniel's and Sara's cases from the previous chapter can be mentioned here to illustrate the above point. Daniel may have concern for safety and Sara may disapprove of the unjust system in which she works. But such attitudes do not by themselves settle appraisability for particular actions. For Daniel to be blameworthy, his driving would need to manifest negligence, indifference, disregard, or some other defective evaluative orientation. If his conduct does not manifest such a defect, blame is not fitting merely because harm occurred, even if the harm is morally serious or wrongful. Sara's case shows the same point differently. Even if Sara recognizes the injustice of the bureaucratic system, her constrained participation would need to manifest

endorsement of the unjust policy, indifference to those harmed by it, or some other blame-grounding orientation. These cases show that evaluative attitudes must be carried through in the relevant action before praise or blame becomes appropriate.

This also supports the claim that manifestation is not just a requirement that connects evaluative attitudes to action. In some cases, the quality of manifestation may be part of what makes agents praiseworthy or blameworthy. Two agents may have the same evaluative attitude about a matter, and both may act from that attitude. But one may show greater care, judgment or effectiveness in carrying that attitude into action, while the other may show carelessness, or may not be effective in enacting the attitude in action. In such a situation, the difference in appraisal does not come from a difference in the underlying attitude alone because the underlying attitude is the same. The difference comes from the agential skills involved in putting one's values into action effectively.

Additionally, this use of Daniel and Sara should be distinguished from the earlier discussion of deliberative control. In the responsibility section, deliberative control was introduced to explain whether an action proceeds through the agent's own practical guidance and therefore whether it is attributable to her. Here, the question is different. Once an action is treated as attributable, manifestation asks a further question: what does that action reveal about the agent's evaluative standpoint? An action may proceed through the agent's control and awareness, and so be attributable to her, without revealing the kind of positive or negative quality that makes the agent appraisable.

Manifestation, then, should be understood as a condition of appraisability in the narrower sense used in this thesis.¹³ The agent is not praised or blamed for manifestation as such. Rather, she is praised or blamed for the action because of what the action manifests about her exercise of agency. If the action manifests care, moral attentiveness, conscientiousness, integrity, or appropriate responsiveness to what matters, praise may be fitting. If it manifests indifference, negligence, selfish disregard, or failure to respond to what the agent could and should have recognized, blame may be fitting. But where the action reveals neither a morally positive nor a morally defective quality, responsibility may remain without either praise or blame being fitting. This is why manifestation is needed as a further condition. Responsibility establishes that the action is attributable to the agent, but appraisability requires that the attributable action reveal something morally significant about the agent's evaluative standpoint.

4.4 Completing the Asymmetry

The account developed in this chapter helps complete the asymmetry between moral responsibility and moral appraisability. Moral responsibility and moral appraisability have different conditions. Moral responsibility requires that the action be attributable to the agent as an exercise of her agency, and this relation obtains through the agent's required control and relevant awareness. Moral appraisability, understood as praiseworthiness or blameworthiness, by contrast, requires something further: the agent's evaluative attitudes must be manifested in the action in a way that makes praise or blame fitting. An action may be attributable to the agent through control and

¹³ To mention again, I do not deny that an attributable action may still be morally assessable in a broader sense, and that the assessment may sometimes be neutral, mixed, or merely lukewarm. My point is that, when moral appraisability is understood as praiseworthiness or blameworthiness, an attributable action must manifest something morally significant about the agent's evaluative standpoint.

awareness, while still failing to manifest a morally positive or defective evaluative standpoint. In such a case, the agent may be morally responsible for the action without being praiseworthy or blameworthy for it. This explains why moral responsibility is not sufficient for moral appraisability. Therefore, this account explains why the intuitive cases discussed in the previous chapter are possible.

At the same time, it is also plausible to infer from the agent-for-action structure of appraisal that moral responsibility is necessary for moral appraisability. If an agent is to be praised or blamed for a particular action, that action must first count as hers in the responsibility-relevant sense. If the action is not attributable to her as an exercise of her agency, then it cannot properly serve as the basis for praising or blaming her. The appraisability conditions therefore presuppose responsibility. That is, the agent's evaluative standpoint must be manifested in an action that is already attributable to her.

Now, we can apply the upshot of this distinction, while keeping the asymmetry between moral responsibility and moral appraisability in mind, to the Rwanda case introduced in **Chapter One** to illustrate how this account distinguishes different kinds of agents within the same broad context of wrongdoing. By looking at the conditions of both responsibility and appraisability, we can say that some agents may be morally responsible as well as morally appraisable. For example, eager or ideologically committed participants may be responsible where their actions are attributable to them through control and awareness; they may also be blameworthy when their conduct manifests endorsement of the wrongdoing, hatred, indifference, or disregard toward the victims. In such cases, responsibility and appraisability come together. Other agents may be morally responsible while their appraisability is more complicated, or in some cases they may be responsible without being blameworthy. In the Rwanda case, this happens with the reluctant or

fearful participants. As I have argued in the responsibility section, the fact that an agent acts under pressure or within constrained options does not always undermine responsibility. A constrained choice may still proceed through the agent's agency and awareness. When this happens, the action remains attributable to the agent. But this does not settle whether or not the agent is praiseworthy or blameworthy. For appraisal asks: what does that choice manifest?

In order to answer the above question, we can consider two fearful participants. The first person participates only because she is under a threat that leaves her with no realistic opportunity for resistance, or escape. Even though her action is attributable to her as it proceeds through her agency under constraint, the action does not manifest endorsement or commitment. She is therefore responsible for the action without being blameworthy for it. The other person is also a participant who acts out of fear. But unlike the first person, she has some opportunity to resist, delay, or reduce harm. She fails to take those opportunities. The point is therefore not that fear or reluctance automatically removes blame. Rather, fear and constraint affect what the action manifests, and so they affect whether blame is fitting and, if it is fitting, how severe it should be.

Similarly, rescuers or refusers may be both responsible and praiseworthy when their actions are attributable to them and manifest courage, care, concern, or responsiveness to the victims' moral standing. Their actions are not merely causally connected to them; they issue from their agency and reveal a morally positive evaluative standpoint. The Rwanda case therefore shows that my account does not deny that responsible agents can be appraisable. It instead shows that moral responsibility is necessary for moral appraisal, even though it is not sufficient.

4.5 Conclusion

This chapter has provided the constructive account needed to complete the asymmetry defended in this thesis. I have argued that moral responsibility and moral appraisability can be distinguished since they are concerned with different questions. The significance of this distinction is that it allows us to recognize a morally meaningful relation between an agent and an action without immediately moving to praise or blame. In this way, the chapter clarifies how moral responsibility can be necessary for moral appraisability without being sufficient for it. In the final chapter of this thesis, I will draw this argument together and will emphasize the broader significance of distinguishing responsibility from appraisability.

Chapter Five

Conclusion

The central question of this thesis has been whether moral responsibility and moral appraisability stand in a biconditional relation. According to this relation, an agent is morally responsible for a particular action if and only if she is morally appraisable for it. I have argued that understanding the relation between moral responsibility and moral appraisability in this way risks collapsing the two different questions: whether an action is attributable to the agent as an exercise of her agency, and whether the agent is praiseworthy or blameworthy for that action. These two moral statuses, I think, need to be understood separately. For this reason, I have argued that the relation between responsibility and appraisability is asymmetric rather than biconditional. That is, moral responsibility is necessary for moral appraisability, but it is not sufficient for it.

5.1 The Argument So Far and Its Scope

The challenge that I have presented to the *Biconditional Assumption* is far-reaching. This assumption is not just an intuitive pattern that reflects our everyday moral judgment. It also appears within the literature on moral responsibility. As I have discussed in *Chapter Two*, the *Desert-Entailing* views make the connection convincing by explaining responsibility in relation to deserved praise or blame. On the *Desert-Entailing* views, an agent is morally responsible when she satisfies the conditions that make praise or blame deserved rather than unfair. These conditions include control, understanding, fair opportunity, avoidability, ownership, and the agent's relation to the action. These conditions determine whether the action is attributed to the agent in a desert-satisfying sense. If an agent lacked control, acted in unavoidable ignorance, or could not reasonably have been expected to avoid the action, then blame would be unfair. This is because the action would not be connected to the agent's agency in the way desert requires. By contrast,

when these excusing conditions are absent, the agent's action is connected to her agency and praise or blame would be deserved. On the **Quality-of-Will** views, moral responsibility is explained by what an action reveals about the agent's will, such as goodwill, ill will, concern, indifference, or disregard. An agent is responsible when the action is expressive of the attitudes that constitute her will. Since praise and blame are fitting responses to good or defective qualities of will, this route also makes responsibility and appraisability appear closely connected.

Even though they differ in what they take to be central, both approaches share a common ground. Each explains moral responsibility in a way that places the agent within the space of moral appraisal. These views make the biconditional relation between responsibility and appraisability appear attractive. The motivation behind these views, particularly the intention of explaining how responsibility remains morally significant, makes the challenge against the biconditional important.

In **Chapter Three**, I have challenged the **Biconditional Assumption** by arguing that responsibility is not sufficient for appraisability. The cases of moral luck and structural injustice were used to show that an agent may be morally responsible for a particular action even though she is not praiseworthy or blameworthy for doing it. By providing counterexamples to the **Biconditional Assumption**, these cases reveal the gap between an action's being attributable to an agent and the agent's being praiseworthy or blameworthy for it.

The constructive account developed in **Chapter Four** explains why and how it is possible for an agent to be morally responsible for an action without being praiseworthy or blameworthy for doing it. On my view, moral responsibility concerns whether an action is attributable to an agent as an exercise of her agency. This account also explains why responsibility continues to

matter without being cashed out in appraisability terms. I have explained that responsibility matters because attributability marks a substantial agential connection between agents and actions. It identifies the action as the agent's own in the responsibility required sense. Control and awareness are the conditions that need to be present for this relation to obtain. Hence, responsibility continues to matter morally even though praise or blame is absent.

Although my account of moral responsibility is stronger than mere causal responsibility, it is still distinct from moral appraisability. An action may be attributable to an agent without the agent being praiseworthy or blameworthy. Moral appraisability requires something further: the agent's evaluative attitudes must be expressed and enacted in deliberation, choice, and action in a way that makes praise or blame fitting. It concerns not merely whether the action is hers but how the action reflects her moral standpoint: what she noticed, valued, ignored, endorsed, rejected, or failed to take seriously.

5.2 Looking Back at the *Desert-Entailing* and the *Quality-of-Will* views

This thesis so far has defended the asymmetric relation between responsibility and appraisability by rejecting the *Biconditional Assumption*. Since *Chapter Two* discussed the *Desert-Entailing* and the *Quality-of-Will* views as two ways to motivate the assumption and my thesis rejects that assumption, it is important to clarify what my view implies for them.

The asymmetric relation developed in this thesis can explain what rejecting the *Biconditional Assumption* implies for the *Desert-Entailing* and the *Quality-of-Will* routes. My view does not claim that either route is wholly mistaken because each route, I think, identifies something important. The *Desert-Entailing* route rightly emphasizes the importance of desert, fairness, and the conditions under which praise or blame would be deserved. The *Quality-of-Will*

route rightly emphasizes the importance of will, regard, concern, and disregard. My objection is not to these concerns themselves, but to the assumption that responsibility and appraisability share a biconditional relation.

My asymmetric account implies that the ***Desert-Entailing*** route needs revision in how it explains the significance of responsibility through deserved praise or blame. The ***Desert-Entailing*** route can still be explained with desert as central, but responsibility needs to be understood first in terms of attributability. My account preserves the desert-entailing theorists' concern with desert but locates the concern at the level of appraisability. That means that desert and fairness mostly matter when we ask whether praise or blame would be deserved for an action that is already attributable to the agent. In ***Chapter Four***, I have explained when an action is attributable to the agent so that the agent is responsible. In this way the desert-entailing theorist can also explain cases like Daniel without having to choose between saying that Daniel is merely causally involved or saying that blame is deserved. Daniel can be responsible because the action is attributable to him, while Daniel is not blameworthy because his action does not satisfy the further conditions for appraisability. The ***Quality-of-Will*** route would have to make a parallel revision. It would have to give up the assumption that every responsible action must already reveal a morally good or defective quality of will. The benefit is that the quality-of-will theorist can preserve the importance of will, regard, and concern through my conception of appraisability without forcing every responsible action into the categories of praise or blame. In this way, some actions can be read as genuinely attributable to the agent without revealing the kind of evaluative standpoint that makes appraisal fitting. In **Section 5.4**, I will show why understanding morally complicated cases through this ***Biconditional Assumption*** generally creates a moral dilemma, whereas the asymmetric relation suggests a better solution.

The advantage of my account, then, is that it gives both routes a way to preserve what they care about while avoiding the ***Biconditional Assumption***. Responsibility remains morally significant because it marks the agent's agential relation to what happened. Appraisability remains important because it concerns what that attributable action reveals about the agent's evaluative standpoint. Separating these two allows responsibility to matter in its own right, without reducing it either to mere causal involvement or to praise and blame.

5.3 Responsibility without Appraisability: the Significance of the Asymmetric Relation

One important upshot of the asymmetric account is that moral responsibility can remain meaningful even when praise or blame is absent. If being responsible does not always make an agent praiseworthy or blameworthy, then why does responsibility still matter? My answer is that responsibility matters because it identifies a substantive agential connection between an agent and an action. It marks the action as attributable to the agent as an exercise of her agency. Responsibility therefore does not merely record that the agent caused something to happen, nor does it already evaluate the agent as praiseworthy or blameworthy. When an action counts as hers in this sense, the agent may have special reasons to acknowledge the event or bear some costs. These reasons do not make the agent eligible for praise or blame.

Recall Daniel's case from ***Chapter Three***. Daniel is involved in bringing about the harm, but the case is structured so that the harm does not reveal a blameworthy evaluative standpoint in the way required for moral appraisability. Even if Daniel is not blameworthy for the resulting harm, the harm remains connected to his agency in a way that distinguishes him from a bystander or someone whose body was moved by force. Daniel knowingly drives in heavy rain and exercises control over what he does. That connection may give him reason to acknowledge what happened,

explain his role, respond to the person affected, or bear some costs. These responses are not grounded in blameworthiness, but they are also not grounded in mere causal involvement.

This is why responsibility can have *normative residue* without making agents praiseworthy or blameworthy. By *normative residue*, I mean the moral significance that remains when an action is attributable to an agent, even though praise or blame is not present. The agent may still have reasons for acknowledgment, repair, or practical response because the action belongs to her agency. The asymmetric relation, then, shows that framing the relation between responsibility and appraisability in this manner preserves the moral significance of responsibility without forcing every responsible action into the categories of praise or blame.

The asymmetry defended in this thesis is important because moral responsibility and moral appraisability are considered foundational notions in both moral philosophy and ordinary moral judgment. In our everyday moral judgment, we ask whether an action is related to the agent in the morally relevant sense so that we can ask: is the action really the agent's? And, is the agent praiseworthy or blameworthy for performing the action? These questions come together even though they are not the same. In this thesis, my attempt has been to show that they are different and concerned with different questions. I have tried to prove that by clarifying the distinction between moral responsibility and moral appraisability. While moral responsibility concerns the agential connection and asks whether an action is attributable to the agent as an exercise of her agency, moral appraisability asks whether that attributable action reveals the agent's evaluative attitudes in a way that makes the agent praiseworthy or blameworthy. Understanding the separation makes it possible to assert that an agent can be morally responsible without being praiseworthy or blameworthy.

This conceptual clarity helps us in moral assessment. If these questions are collapsed, then moral assessment fails to distinguish between the question of whether an action is attributable to the agent and the question of whether that action reveals her evaluative standpoint in a way that makes praise or blame fitting. This leaves us with one of two unsatisfactory results: either denying that the agent is responsible whenever praise or blame is unfitting, or treating praise or blame as fitting whenever responsibility is present. The framework defended in this thesis avoids both mistakes. It allows us to recognize a responsibility-relevant relation between agents and actions while leaving open the further evaluative question of whether the agent is praiseworthy or blameworthy. In this way, the account makes moral assessment more precise by separating the question of agential connection from the question of praiseworthiness and blameworthiness.

5.4 Implications for Morally Complicated Cases

The most noteworthy payoff of my framework is that it offers a way to approach morally complicated cases without forcing them into a single category. Cases involving structural injustice and moral luck often create a moral dilemma in which we have to move either toward blame or toward the denial of responsibility. This section will explain what these cases look like, how the biconditional relation would read them, and how the moral dilemma arises from that reading. By referring back to my *Chapter Three* examples of Sara and Daniel, I will give an alternative interpretation of these cases, which I believe can help us address the moral dilemma.

Structural Injustice: structural injustice arises when harm or disadvantage is produced by social arrangements and institutional practices where no single individual is identified as responsible for creating the arrangements or enhancing the practices. The impact of these social arrangements and practices might be clearly visible, but it is not easy to locate one direct wrongdoer. We are aware of this injustice. Workers around the world are placed under unfair

constraints by economic conditions, and they get lower wages compared to the work they do. Unequal power distribution, social roles, and limited opportunities are the results of structural injustice. Recall the Sara case from **Chapter Three**. Sara is working in an unjust bureaucratic system. She does not endorse the system as she is aware of its unjust policies. She also knows that she cannot change the system with her limited power even if she tries to do so. Also, she has to run her family with the income. So, she continues to perform her role in that system.

From one perspective, we can say that within structural injustice, Sara is not blameworthy. The reason is that her action takes place within a background of pressure, limited options, institutional expectations, and constrained practical circumstances. These factors are important for assessment because they affect what her action manifests about her evaluative attitudes. Sara's continued participation does not express endorsement of the system, instead, it reflects the difficult position in which she is placed by the overall structure.

However, it would actually be unfair to blame her in the same way as someone who willingly endorses the unjust system because Sara's social surroundings and her economic status have limited her agency. Her action might not express endorsement of the injustice; instead, it may reflect her difficult position in which she is now because of the overall structure. However, the other perspective gives a different verdict. Sara is not just a bystander, nor is she simply a passive victim of the unjust system. She continues to act in that system and contributes to some extent to the functioning of the unjust system, even if only on a very minimal scale. Her action, then, cannot be understood as something just happening around her. It is attributable to her as an exercise of her agency even if the exercise of her agency is limited and constrained by the unjust system. This perspective, therefore, gives a different result. Here, Sara is morally responsible for what she does.

If moral responsibility and moral appraisability share a biconditional relation, then we have to accept one of the two unsatisfactory judgements. Either we say that Sara is responsible, and therefore she is also blameworthy for what she does. Or we have to say that she is not blameworthy for participating in an unjust system, and she is therefore not responsible. Accepting the first one would be unfair to Sara because Sara's constrained participation would be treated as if it manifested the same blame-grounding attitude as that manifested by someone who willingly endorses the unjust system. The second option is also problematic because it erases Sara's agency, which was constrained but not totally extinguished.

My asymmetric account proposes a way out of this dilemma. It allows us to understand and separate two questions. First, is Sara's action attributable to her as an exercise of her agency? Second, does that attributable action manifest the evaluative standpoint required for blame? The answer to the first question can be yes, whereas the answer to the second question is no. Sara may be responsible because her action may be connected to her in the way that preserves the agential connection required for moral responsibility. But this does not settle whether or not she is praiseworthy or blameworthy. Whether she is blameworthy or not depends on what her constrained participation expresses about her evaluative attitudes and how those attitudes are manifested. If her action manifests endorsement, then blame would be appropriate. If the structural constraint affects what her action manifests about her evaluative attitudes, then she would not be blameworthy.

Reading the above case, within my framework, leaves room for analyzing different cases differently even though both are cases of structural injustice. This matters significantly because structurally complicated cases are common and they show a problem we often face when we try to judge agents who act under pressure and limited options created by social or institutional

systems. Some agents may be both responsible and blameworthy in this system. Unlike Sara, a person who endorses the system may manifest hatred and endorsement is responsible as well as blameworthy. Therefore, we can say that responsibility does not settle appraisability. In structurally complicated cases, we are better able to reach a decision when we understand and separate moral responsibility and moral appraisability.

Moral Luck: moral luck provides another case in which the asymmetric relation between moral responsibility and moral appraisability proves valuable. Moral luck related cases occurs when factors outside agents' control affect how agents' actions are morally assessed. For instance, two drivers drive their cars while they are both too sleepy and get behind the wheel when they should not. One of them makes it home without hurting himself or anyone; the other hurts someone. Now, how they are assessed for their actions depends to some extent on something other than what they do. Here, luck makes the difference and enters the assessment process. We are familiar with this type of case in our ordinary life. There are situations in which we have very different judgments about agents, and it looks like we judge someone blameworthy. We think people are excused from blame when they cannot control what they do. We feel justified in doing this. But the example shows that the difference between the person we judge blameworthy and the person we do not judge blameworthy is sometimes just about a matter of luck over which neither person has control.

We can recall the Daniel case to present this difficulty.¹ Daniel drives in heavy rain. He is a careful driver who does not act with disregard for others. Still, because of factors outside his control, a child suddenly appears and the child is harmed. Focusing on the luck issue, we might

¹ To be mentioned that Daniel's case is different from the sleepy drivers' case because Daniel is not presented as a negligent agent.

say that Daniel is not blameworthy and blaming him would be unfair. For the outcome is influenced by an outside factor, *i.e.*, the child's sudden appearance. Another equally careful driver might have driven in the same way and harmed no one simply because no one suddenly appeared on his way. This suggests that the unlucky outcome alone should not make Daniel blameworthy.

However, Daniel's agency has not disappeared either. He is not a bystander, nor is he merely causally connected to the action. The action did not merely occur through his activity; instead, he knowingly undertook his decision to drive. That is, Daniel exercised his agency as the agential connection was present. For that reason, Daniel may have special reasons to acknowledge, express regret, assist the child, or bear some of the costs. However, these responses are not grounded in blameworthiness. They are grounded in the fact that Daniel is a moral agent who knowingly exercises his agency. I understand these responses as *normative residues*. I think that these normative residues remain when an agent is morally responsible and that do not require praise or blame. They refer to the moral significance that remains when an action is attributable to an agent, even when praise or blame is not present. Agent-regret is one important form of such residue. A person may appropriately feel a distinctive regret when something harmful happens through her own agency even though she is not at fault.² But this agent-regret does not show that Daniel is blameworthy; it just shows that he is not related to the action as a bystander but is related through an agential connection.

² Bernard Williams, "Moral Luck," in *Moral Luck: Philosophical Papers 1973–1980* (Cambridge: Cambridge University Press, 1981), 28, <https://doi.org/10.1017/CBO9781139165860.003>.

The dilemma that appears in Sara's case also appears in Daniel's case when Daniel is understood through the ***Biconditional Assumption***. The difficulty is that Daniel's case contains both agency and luck. On the one hand, Daniel's case is not a case of pure brute luck. Brute luck refers to what happens to a person without any relevant choice on her part, as when someone is struck by lightning or harmed by a natural disaster. Daniel's case is different because he knowingly chooses to drive in heavy rain. His action therefore involves a chosen risk, and this chosen risk helps explain why the harm remains connected to his agency.

On the other hand, Daniel's case is not a case in which the harmful outcome is fully under his control. The child's sudden appearance, the timing of the event, and the fact that harm actually occurs are partly matters of luck. Daniel does not choose to harm the child, nor does the bad outcome by itself show that he acted with negligence, indifference, or disregard. For this reason, the lucky or unlucky outcome should not automatically make him blameworthy.

The ***Biconditional Assumption*** has difficulty with either decision. If the case is treated only as a case of chosen risk, then Daniel's responsibility may seem to lead directly to blame. If the case is treated only as a case of luck, then blame may seem unfitting only because Daniel is not responsible at all. Both interpretations suggest something misleading. Daniel is responsible because the harm is connected to his agency through his knowingly risk-imposing action. But he is not blameworthy because the action does not manifest a defective evaluative standpoint. The advantage of my account is that it preserves the moral significance of Daniel's agency without converting that significance into blame. It also explains why we should not infer moral fault from harm. It, again, explains why Daniel may have reasons for regret, acknowledgment, assistance or repair, while also explaining why these reasons do not amount to blameworthiness. Moral luck therefore supports the need for asymmetry. It shows that we need a middle position between

blameworthiness and mere causal involvement: responsibility without fault, or *normative residue* without blame.

5.5 Concluding Remarks

This thesis started with the Rwanda case, a morally complicated case that might be read from different perspectives, and from those perspectives different questions concerning responsibility and appraisability arise. Exploring those different perspectives has aided us in realizing that responsibility and appraisability are concerned with different questions. This does not mean that I deny that morally responsible agents can be morally appraisable. Rather, my claim is that the relation between responsibility and appraisability should not be understood as one in which responsibility is sufficient for appraisability. I have presented morally complicated cases in which agents are morally responsible for their actions but are not appraisable for them. However, I think that being responsible is necessary for being appraisable; thus, I have suggested an asymmetric relation between them.

This asymmetric relation is the main contribution of my thesis. By separating responsibility from appraisability, I think that we can gain a clearer way to think about moral agency, blame, praise, excuse, luck, and constraint. Additionally, this relation preserves the thought that responsible agency matters, while resisting the assumption that every responsible action must already make the agent praiseworthy or blameworthy. Moral responsibility, on this view, marks the agent's relation to what she has done. Moral appraisability asks what that action reveals about her evaluative standpoint. Keeping these two questions apart gives us a more careful and more accurate way of understanding moral responsibility without losing its significance.

—

Bibliography

- Arpaly, Nomy. *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford: Oxford University Press, 2003.
- Bagnoli, Carla. "Claiming Responsibility for Action Under Duress." *Ethical Theory and Moral Practice* 21, no. 4 (2018): 851–868. <https://doi.org/10.1007/s10677-018-9904-8>.
- Campbell, Robert. "Duress and Responsibility for Action." *Journal of Applied Philosophy* 1, no. 1 (1984): 133–140.
- Frankfurt, Harry G. "Alternative Possibilities and Moral Responsibility." *The Journal of Philosophy* 66, no. 23 (1969): 829–839. <https://doi.org/10.2307/2023833>.
- Grant, James. "Moral Worth and Moral Belief." *Ethics* 133, no. 2 (2023): 216–230. <https://doi.org/10.1086/722120>.
- Gunnemyr, Mattias, and Caroline Torpe Touborg. "You Just Didn't Care Enough: Quality of Will, Causation, and Blameworthiness for Actions, Omissions, and Outcomes." *Journal of Ethics and Social Philosophy* 24, no. 1 (2023): 1–35. <https://doi.org/10.26556/jesp.v24i.1696>.
- Haji, Ishtiyaque. *Moral Appraisability: Puzzles, Proposals, and Perplexities*. New York: Oxford University Press, 1998.
- King, Matt. "Moral Responsibility and Merit," *Journal of Ethics and Social Philosophy* 6, no. 2 (2012): 1–18.
- King, Matt. "Two Faces of Desert." *Philosophical Studies* 169, no. 3 (2014): 401–24. <https://doi.org/10.1007/s11098-013-0188-5>.
- Madigan, Taylor. "Two Dimensions of Responsibility: Quality and Competence of Will." *Journal of the American Philosophical Association* 10, no. 2 (2024): 281–294. <https://doi.org/10.1017/apa.2023.2>.

- Mamdani, Mahmood. *When Victims Become Killers: Colonialism, Nativism, and the Genocide in Rwanda*. Princeton, NJ: Princeton University Press, 2001.
- McMahan, Jeff. “The Basis of Moral Liability to Defensive Killing.” *Philosophical Issues* 15 (2005): 386–405. <https://doi.org/10.1111/j.1533-6077.2005.00073.x>.
- Nagel, Thomas. “Moral Luck.” *Proceedings of the Aristotelian Society, Supplementary Volumes* 50 (1976): 137–151.
- Nelkin, Dana Kay. “Accountability and Desert.” *The Journal of Ethics* 20, nos.1-3 (2016): 173–189. <https://doi.org/10.1007/s10892-016-9230-0>;
- Nelkin, Dana Kay. “Desert, Fairness, and Resentment.” *Philosophical Explorations* 16, no.2 (2013): 117–132. <https://doi.org/10.1080/13869795.2013.787438> .
- Nelkin, Dana Kay. “Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness.” *Nous* 50, no. 2 (2016): 356–378. <https://doi.org/10.1111/nous.12079>.
- Pereboom, Derk. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press, 2014.
- Rosen, Gideon. “Culpability and Ignorance.” *Proceedings of the Aristotelian Society* 103, no.1 (2003): 61–84.
- Sartorio, Carolina. “Causation and Responsibility.” *Philosophy Compass* 2, no. 5 (2007): 749–765. <https://doi.org/10.1111/j.1747-9991.2007.00097.x>.
- Shoemaker, David. “Qualities of Will.” *Social Philosophy and Policy* 30, nos. 1–2 (2013): 95–120.
- Smith, Angela M. “Blame and Holding responsible.” In *The Oxford Handbook of Moral Responsibility*, ed. Dana Kay Nelkin and Derk Pereboom, 269–286. New York: Oxford University Press, 2022. <https://doi.org/10.1093/oxford/9780190679309.013.15>.
- Smith, Holly. “Culpable Ignorance.” *The Philosophical Review* 92, no. 4 (1983): 543–71.

- Strawson, P. F. "Freedom and Resentment." *Proceedings of the British Academy* 48 (1962): 187-211.
- Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, 1994.
- Williams, Bernard. "Moral Luck." In *Moral Luck: Philosophical Papers 1973–1980*. Cambridge: Cambridge University Press, 1981. <https://doi.org/10.1017/CBO9781139165860.003>
- Wolf, Susan. *Freedom Within Reason*. Oxford: Oxford University Press, 1990.
- Young, Iris Marion. *Responsibility for Justice*. New York: Oxford University Press, 2011.
- Zheng, Robin. "What Is My Role in Changing the System? A New Model of Responsibility for Structural Injustice." *Ethical Theory and Moral Practice* 21 (2018): 869–885. <https://doi.org/10.1007/s10677-018-9892-8>.
- Zimmerman, Michael. "Moral Responsibility and Ignorance." *Ethics* 107, no. 3 (1997): 410–26.