

MORAL WRONGNESS AND REACTIVE ATTITUDES

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

WENWEN FAN

Dr. Peter Vallentyne, Dissertation Supervisor

MAY 2014

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

MORAL WRONGNESS AND REACTIVE ATTITUDES

presented by Wenwen Fan,

a candidate for the degree of doctor of philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Peter Vallentyne

Professor Robert Johnson

Professor Peter Markie

Professor Paul Litton

Professor Matthew McGrath

For Chang, my Mom and Dad

ACKNOWLEDGMENTS

I would like to express the deepest appreciation to my adviser, Professor Peter Vallentyne, who guided and supported me through my whole dissertation writing process. He helped me focused and disciplined in my work, offered prompt and detailed comments, and encouraged me to keep working when I felt disappointed at my work. Without his guidance and persistent help, this dissertation would not have been possible.

I would also like to thank my committee members, Professor Robert Johnson, Professor Peter Markie, Professor Paul Litton, and Professor Matthew McGrath. They each discussed with me about my dissertation and offered insightful comments. I benefited enormously from their help and felt blessed to work with such intelligent and perceptive professionals.

I am also grateful to Tom Reynolds, Richard Lauer, Jonah Goldwater for helpful discussions. In addition, a thank you to the audience at the Rocky Mountain Congress in 2013, where I presented part of my work. The comments and objections helped me refined my view and alerted me to issues I needed to address.

Last but not least, I am grateful to my loving parents and supportive husband. They always supported me when I had doubts, and they always had confidence about my abilities. Their support is truly invaluable for me.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
ABSTRACT	v
Ch. 1 Introduction	1
1. What Are Moral Wrongness and Reactive Attitude	2
2. How Might Moral Wrongness and Reactive Attitudes Be Connected	6
3. Philosophical Significance	7
4. The Features of Reactive Attitudes	8
5. Main Positions	11
6. Conclusion	16
Ch. 2 The Naturalist Reactive Attitude Approach	18
1. The Naturalist Reactive Attitude Approach	18
2. Motivations for the Naturalist Reactive Attitude Approach	22
3. Problems of the Naturalist Reactive Attitude Approach	23
4. Conclusion	32
Ch. 3 The Non-Naturalist Reactive Attitude Approach	33
1. The Non-Naturalist Reactive Attitude Approach	33
2. Motivations for the Non-Naturalist Reactive Attitude Approach	37
3. Replies to Critics	38
4. Conclusion	45
Ch. 4 The Reactive Attitude View	47
1. The Reactive Attitude View	47
2. Motivations for The Reactive Attitude View	56
3. Problems of The Reactive Attitude View	57

4. Conclusion	84
Ch. 5 The Interdependence View and The Independence View.....	85
1. Against The Interdependence View	85
2. Against The Independence View	90
Ch. 6 The Standard View.....	97
1. The Standard View	97
Ch. 7 Conclusion.....	127
BIBLIOGRAPHY.....	136
VITA.....	142

ABSTRACT

In my dissertation, I examine the relationship between moral wrongness and negative reactive attitudes. In particular, I inquire (1) whether moral wrongness is conceptually connected to the *empirical disposition* to hold certain negative reactive attitudes or the *normative status* (e.g., appropriateness) of holding certain reactive attitudes, and (2) if so, which is conceptually prior. I argue that moral wrongness is not conceptually connected to the empirical disposition to hold certain negative reactive attitudes, but moral wrongness is conceptually connected to the fittingness of holding certain negative reactive attitudes towards the agent who inexcusably acts wrongly. I further argue that wrongness is conceptually prior to the fittingness of holding certain reactive attitudes, but not vice versa.

Ch. 1 Introduction

The main question this dissertation addresses is the relationship between moral wrongness and reactive attitudes. This question can be asked in several ways: First, are moral wrongness and reactive attitudes ever related? If they are not, why is that? If they are, are they necessarily related or contingently related, and is moral wrongness related to every type of reactive attitude or is it merely related to certain types of reactive attitudes (e.g., blame, guilt)?

Second, is the concept of moral wrongness prior to the concept of reactive attitude, or vice-versa, or are these concepts interdependent? If concept A is prior to concept B, then we can analyze B by A without circularity, but we cannot analyze A by B without circularity. In other words, if concept A is prior to concept B, then we must understand concept B in terms of concept A, but not vice versa. Note that the issue of conceptual priority is not about which concept is in fact acquired before the other concept, but which concept is more basic or more primitive in terms of meaning. Regarding the project of this dissertation, the question can be: If the concept of moral wrongness is prior to the concept of reactive attitude, then does moral wrongness, at least partially, determine or ground the empirical disposition, or the normative status of our having (or expressing) certain reactive attitudes? Alternatively, if the concept of reactive attitude is prior to the concept of moral wrongness, then does the normative status of our having (or expressing) certain reactive attitudes determine moral wrongness?

The focus of this dissertation is the second sort of the questions presented above, namely, the metaethical question about the relationship between the concept of moral wrongness and the concept of reactive attitude. As the discussion moves on, the

2dissertation will focus on particular reactive attitudes such as blame, indignation, resentment, and guilt. Thus, relevant to the approach discussed, the precise question will become the relationship between the concepts of moral wrongness and blame, or the relationship between the concept of moral wrongness and the concepts of resentment, indignation, and guilt.

1. What Are Moral Wrongness and Reactive Attitude

Since my core question is the relationship between moral wrongness and reactive attitudes, I need to explain how I understand these terms. Moral wrongness is the violation of moral requirements. An action is morally wrong if and only if it violates at least one moral requirement. A morally wrong action is also morally impermissible. In contrast, a morally right action is morally permissible: it violates no moral requirement. So, a morally right action is either morally obligatory or morally optional.

Reactive attitudes, as a first approximation, are certain kinds of emotion that we have towards a person in response to the manner she behaves. They are roughly emotions we naturally have when we participate in human relationships. In contrast, we hold objective attitudes towards other people when we regard them as objects (e.g., subjects of human research). For instance, resentment is a reactive attitude. Nicole may harbor resentment towards her boss, Mark, because he passes her over for promotion in order to promote his relative. Nicole's resentment may be naturally explained by saying that Mark treats her unfairly and she resents being treated unfairly. Besides resentment, it is commonly accepted that gratitude, indignation, and guilt are typical reactive attitudes (Butler, 1726; Strawson, 1962; Wallace, 1994).

Reactive attitudes can be negative, positive, or neutral. We normally harbor negative reactive attitudes when we feel *bad* towards someone regarding her action, intention, character, etc., while we normally harbor positive reactive attitudes when we feel *good* towards someone regarding her action, intention, character, etc. For instance, if Sue helps Brian with his project, it is natural for Brian to feel gratitude towards her because of her help. We normally harbor neutral reactive attitudes when we feel *indifferent* towards someone regarding her action, intention, character, etc. For instance, if Lynda scratches her head in class, her professor normally has no particular feelings (approximately disinterested) towards her regarding her action.

Unfortunately, no unique definition of reactive attitudes is commonly accepted. Historically, reactive attitudes and similar notions have been discussed in the names of ‘sentiment,’ ‘moral sentiment,’ ‘reactive feelings,’ and ‘reactive emotion’ (Hume, 1751; Mill, 1861; Wallace, 1994). The core idea is the emotions human beings have towards themselves or others regarding the target person’s character trait, action, attitude, intention, etc. A wide notion of reactive attitude simply explains the literal meanings of ‘reactive’ and ‘attitude’ and leaves it open whether any belief about (moral or nonmoral) normative behaviors is involved. For example, P. F. Strawson (1962) essentially defines reactive attitudes as “natural human reactions to the good or ill will or indifference of others towards us, as displayed in their attitudes and actions.”¹ In contrast, a narrow conception of reactive attitude may regard the beliefs about normative behaviors to be essential to reactive attitudes. For example, W. J. Wallace (1994) holds that reactive

¹ Strawson, 1962, p. 195. Apart from attitudes towards others (on behalf of oneself), Strawson also talks about attitudes towards oneself (on behalf of others) and attitudes towards others on behalf of others. He seems to think that attitudes towards others (on behalf of oneself) are central reactive attitudes.

attitudes are “the emotions that are constitutively linked to *expectations*, in the sense of holding someone to an expectation,” where an expectation is the agent’s belief about (normative) requirements.² Moreover, a narrow conception may even regard the beliefs about normative behaviors to be exclusively moral. For instance, John Rawls (1971) goes as far as to claim that explaining moral sentiments characteristically invoke *moral* concepts (e.g., the concepts of the right and the good) and associated principles.³

Since philosophers don’t agree with each other about what reactive attitudes mean, it is natural that they don’t agree about which emotions count as reactive attitudes. Yet, they generally agree that admiration, gratitude, resentment, indignation, guilt, and shame are reactive attitudes (Butler, 1774; Strawson, 1962; Wallace, 1994). For this reason, I will focus on these attitudes. More precisely, I will focus on resentment, indignation, guilt, and blame. I will not focus on positive reactive attitudes because attitudes such as admiration and gratitude are more complicated than negative ones: for instance, they do not connect to our normative desires or beliefs as closely as negative ones do (more about this is in Section 3.3). Accordingly, all the reactive attitudes I focus on are negative. Among negative reactive attitudes, shame will not be discussed because it is a more complicated emotion than resentment and guilt, and no authors to discuss use shame to analyze moral wrongness. Furthermore, I will regard blame as a reactive attitude, one that appears in the fashion of resentment, indignation, or guilt. When I discuss blame, I am not talking about a judgment about an action, etc, but a sort of negative feelings about an

² Wallace, 1994, p. 21, my emphasis.

³ Rawls, 1971, pp. 481-482.

action, an agent, or an intention, etc.⁴ For simplicity's sake, I assume cognitivism (the view that moral judgments express beliefs), which enables the talk of resentment and guilt, etc. to connect with beliefs. I will make appropriate adjustments when non-cognitivism needs to be discussed.

For now, I will offer a broad definition of reactive attitude to facilitate the current discussion, and I will examine different positions according to the particular conception of reactive attitude the person discussed adopts. My definition is this: reactive attitudes are *the emotions an agent holds towards another agent or herself in response to what she believes to be the manner in which the target agent exercised her agency, whereas she regards the target agent as a fellow human being rather than a research object*. Some clarifications for my definition are called for. 'The manner,' first and foremost, refers the target agent's way (e.g., level of carefulness, used tools) in which she exercised her agency. Moreover, 'the manner' refers to the target agent's mental state (e.g., intention, thoughts) when she exercised her agency. For instance, if Sally has a compulsive desire to steal a book from a bookstore but manages to refrain from acting on the desire, then we might admire her when we know the fact. As for regarding 'the target agent as a fellow human being rather than a research object,' I intend to capture Strawson's idea that reactive attitudes merely happen among interpersonal relationship, the idea that when in this relationship, we will regard each other as equal beings exercising agency instead of objects to be studied in a lab. I admit that my definition is vague, but that suffices for facilitating the discussion.

⁴ I acknowledge that the notion of blame allows several conceptions: a negative assessment of one's character, a kind of sanction, or an indication of the impairment of two people's relationship. See Scanlon, 2008. I do not claim that my conception of blame is the best or the only correct one.

2. How Might Moral Wrongness and Reactive Attitudes Be Connected

Having introduced the notions of moral wrongness and reactive attitude, I will discuss two ways in which moral wrongness may be connected with reactive attitudes. First, there is the distinction between *having* a reactive attitude and *expressing* it. Moral wrongness may be connected with either. Having an attitude is feeling the relevant sentiment or emotion, while expressing an attitude is acting out the emotion in words or the other kinds of behaviors. For example, I may judge a person who drives his car into the sidewalk and hurts a pedestrian to be blameworthy, but I may not express blame to anyone about the driver. In this case, I hold the blame-sentiment but I may not express blame. My working assumption is to discuss having an attitude instead of expressing an attitude.

Second, there is also the distinction between *being appropriate* to have (or express) an attitude and *having a moral obligation* to have (or express) an attitude. Moral wrongness may be connected with either. If one's holding an attitude is appropriate, then the attitude is fitting, given the action performed. For instance, if Boss passes Employee over for promotion in order to promote his relative, it is appropriate for Employee to feel resentment towards Boss regarding his action, other things being equal. We can say that Employee's feeling resentment is appropriate because it fits Boss's action and intention. In contrast, one's having a moral obligation to hold an attitude means that morality requires the person to hold the attitude, which may have nothing to do with the action performed. In the case just discussed, Employee may have a moral obligation not to feel resentful at Boss because her resentment generates horrible consequences. Hence, it can

be appropriate for an agent to hold an attitude towards another agent regarding an action, while she has no moral obligation to hold that attitude. The opposite can also be true: an agent may have a moral obligation to hold an attitude towards another agent regarding an action because it will have very good consequences, but it is not appropriate for her to hold that attitude. My working assumption is to discuss being appropriate to have an attitude, not having a moral obligation to have an attitude.

Note that the phrase ‘being appropriate to have an attitude’ is ambiguous, for the appropriateness could be regarding the real situation or the situation believed by the reacting person. I will call the sort of appropriateness regarding the real situation *objective appropriateness* and the sort of appropriateness regarding the reacting agent’s believed situation *subjective appropriateness*. Suppose Agent helped Patient, other things equal, but Perceiver mistook him to have harmed Patient and have no excuse. Then it is subjectively appropriate for Perceiver to feel the blame-sentiment towards Agent regarding his action, but it is not objectively appropriate for Perceiver to feel so. My working assumption is to discuss being objectively appropriate to have an attitude, not being subjectively appropriate to have an attitude.

3. Philosophical Significance

The question about the conceptual relationship between moral wrongness and reactive attitudes matters for several reasons. First, this question is relevant to the debate on moral responsibility. Ever since Strawson drew philosophers’ attention to the role of reactive attitudes, these attitudes have been deemed important regarding theoretical questions such as when we are morally responsible for our actions. Since moral wrongness

obviously connects to what we are morally required to do and thus connects to what we are morally responsible for (in the sense of accountability), it is important to have an account about the conceptual relationship between moral wrongness and reactive attitudes. Moreover, since moral wrongness is connected with important notions such as moral permissibility and moral obligation, we can gain more knowledge about these notions if we know more about moral wrongness.

Second, answering this question may shed some light on understanding emotions and moral assessments. On the one hand, emotions cover non-cognitive feelings, cognitive thoughts, and dispositions to act. Since reactive attitudes are a sort of emotions, understanding how such attitudes connect to moral wrongness may help understand what emotions are. On the other hand, some think that moral rightness/wrongness is disconnected from praise/blame. If so, it seems natural to infer that moral wrongness is disconnected from reactive attitudes (since praise and blame, widely speaking, are reactive attitudes). Yet no common opinions have reached on this matter. Hence, studying the relationship between moral wrongness and reactive attitudes is helpful.

4. The Features of Reactive Attitudes

4.1 In What Sense Are Reactive Attitudes Reactive

Before I sketch the main positions, I shall further clarify the notion of reactive attitude. The first purpose of this section is to fully introduce a contrasting notion with reactive attitudes: objective attitudes. The second purpose is to explain the controversy over whether reactive attitudes are moral (i.e., connect with moral desires or beliefs), normative but not necessarily moral, or neither.

As Strawson (1962) points out, reactive attitudes contrast with objective attitudes. Objective attitudes are the emotions characteristic of reactions to objects. For example, when a physicist studies a newly found object, she does not regard the object as an agent or form any human (or, agential) relationship with it, but she takes an objective stance towards it. Similarly, a mother may take a detached stance towards her 5-year-old son when she educates him. She normally feels excited or frustrated because of his son's response, and the excitement and frustration of this kind are objective attitudes. In Strawson's words, adopting objective attitudes essentially is "a consequence of our viewing the agent as incapacitated in some or all respects for ordinary interpersonal relationships."⁵ We can roughly understand what he calls 'ordinary interpersonal relationships' as relationships formed by agents who exercise agency. Accordingly, people have objective attitudes towards others when they perceive the target person either not as a fellow human at all, or as an agent that did not exercise her agency. Thus, professionals like therapists and special education teachers may adopt objective attitudes towards their target people when they work. Objective attitudes are also held when the target agent is not autonomous (e.g., being manipulated), does not possess the relevant knowledge, or does not act voluntarily (e.g., being coerced).

In contrast with objective attitudes, reactive attitudes are taken when the reacting agent regards the target person as a fellow human who exercises her agency. For example, Nicole feels resentful at her boss, Mark, for passing her over for promotion because she

⁵ Strawson, 1962, p.198. In order to illuminate the contrast between objective and reactive attitudes, he seems to postulate that objective attitudes are held against certain agents for a relatively long time because those agents are incapable of being autonomous (e.g., being a schizophrenic). I will interpret objective attitudes more widely, taking them to be held against any agent that does not exercise agency (either temporarily or permanently).

believes that he exercised his agency in doing so. Likewise, Jay feels guilty about not preventing a big man from raping a girl because Jay believes himself to have exercised his agency in walking away from the raping scene.

4.2 Are All Reactive Attitudes Moral, Normative, etc.

Not all agree whether all reactive attitudes characteristically connect with moral desires or beliefs. For example, Wallace (1994, p. 36) argues that reactive attitudes are only moral when they are linked to obligations for which the attitude holder is able to provide *moral* justifications, whereas Rawls (1971, p. 481) claims that all reactive attitudes are characteristically moral because the presence of these attitudes should be explained by *moral* concepts and associated principles. The crux is that a reactive attitude is characteristically moral if and only if explaining its characteristic presence must appeal to moral concepts or moral principles. I will not take a stance on this matter, since I want to accommodate different conceptions of reactive attitude.

Even if not all reactive attitudes are moral, they might still be normative. Many (e.g., Wallace, 1994; Darwall, 2006) maintain that all reactive attitudes connect with normative desires or beliefs. In fact, Strawson seems to agree on this matter, since he connects reactive attitudes with expectations or demands.⁶ It seems truism that reactive attitudes are normative, or at least characteristically so. Indeed, we hold reactive attitudes because we respond to people's actions relative to certain standards, the standards about how people *should* think, act, etc. I may feel guilty about not helping a stranger because I

⁶ Strawson, 1962, p. 200.

think that I should have helped her, and you may feel indignant about not being accepted to a prestigious graduate school because you think that you *should* be accepted.

However, Hume (1739-40) seems to disagree about the normativity of reactive attitudes. According to Hume, reactive attitudes are merely particular pains or pleasures that we naturally feel.⁷ No normative desires or beliefs necessarily connect with such attitudes. This view makes sense if we see reactive attitudes as a kind of secondary quality (e.g., color) that we perceive through our faculties, or more precisely, feelings that are simply imposed on us by people's behaviors. Again, I will not take a stance on whether all reactive attitudes are (characteristically) normative in order to maintain neutral on my understanding of reactive attitudes.

Despite my neutral position, since my project is the relationship between *moral* wrongness and certain negative reactive attitudes (resentment, indignation, guilt, and blame), I mainly explore the situations where deontic assessment is involved. This means that only cases where right and wrong are concerned will be discussed, regardless of whether resentment, indignation, guilt, and blame can be held in non-moral cases.

5. Main Positions

My core question is: What is the relationship between the concept of moral wrongness and particular reactive attitudes? More precisely, the question becomes: What is the relationship between the concept of moral wrongness and the concepts of blame, resentment, guilt, indignation, etc.? Recall that I will focus on negative reactive attitudes

⁷ Hume, 1739-40, 3.1.2.3.

and I regard blame, resentment, guilt, and indignation as the main paradigms. In this section, I will introduce the main positions that I shall examine.

We can develop two main approaches to the question, depending upon their answers to the following question: Is moral wrongness conceptually connected with the *empirical disposition* or the *normative status* (appropriateness, permissibility, or obligation) of our having (or expressing) negative reactive attitudes such as blame, resentment, indignation, and guilt? If one's answer is the empirical disposition, then I will call the approach the *Naturalist Reactive Attitude Approach*. If one's answer is the normative status, then I will call the approach the *Non-Naturalist Reactive Attitude Approach*. If one answers the first question negatively but the second question affirmatively, then I will call the view *Reactive Attitude View*. In the remainder of this chapter, I will sketch these approaches.

5.1 The Naturalist Reactive Attitude Approach

The Naturalist Reactive Attitude Approach claims that moral wrongness is conceptually connected to our empirical disposition to have (or express) core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt) (Hume, 1739-40, 1751; Strawson, 1962?)

This approach is often held by moral naturalists such as Humeans. These people believe that morality is as natural as science such that we can discover connections between morality and the natural phenomena. Along this vein, since it is an empirical matter to have the disposition to feel the blame-sentiment when we react to people's

wrongful actions, it might just as well that such disposition is connected to the concept of moral wrongness.

5.2 The Non-Naturalist Reactive Attitude Approach

The Non-Naturalist Reactive Attitude Approach claims that moral wrongness is conceptually connected to the normative status (appropriateness, moral permissibility, etc.) of our having/expressing core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt) (Mill, 1861; Gibbard, 1990; Darwall, 2006; Skorupski, 2010a, 2010b).

This approach is attractive if one thinks that the normativity of morality is unlikely to be an empirical matter. Thus, although proponents of the Non-Naturalist Reactive Attitude Approach find the practice of having/expressing reactive attitudes important, they hold that merely the normative status of this practice is connected to the concept of moral wrongness.

Among proponents of the Non-Naturalist Reactive Attitude Approach, three views are available. The first view holds that moral wrongness is conceptually prior to the normative status of our having/expressing core negative reactive attitudes, but not vice versa. I call this view the *Standard View*. The second view holds that the normative status of our having/expressing core negative reactive attitudes is conceptually prior to moral wrongness, but not vice versa. I call this view the *Reactive Attitude View*. The third view holds that neither moral wrongness nor the normative status of our having/expressing core negative reactive attitudes is conceptually prior to the other. I call this view the *Independence View*. Let me briefly explain them.

The Standard View

The Standard View claims that moral wrongness is conceptually prior to the normative status of our having/expressing core reactive attitudes (i.e., blame, resentment, indignation, and guilt), but not vice versa (Rawls, 1971; arguably Ross, 1930, 1939; arguably Kant, 1785, 1788).

Holders of the Standard View find the notion of moral wrongness intelligible without any appeal to empirical experiences. They hold that agents normally have the capacity to understand the abstract ideas of moral requirements, moral permissibility, and moral obligation, etc. Thus, the appropriateness, permissibility, or obligation of having (or expressing) reactive attitudes is not required to understand moral wrongness.

Furthermore, they maintain that the essence of reactive attitudes is not merely emotions in human interactions, but also the connection with moral demands or moral requirements. Hence, the appropriateness, permissibility, or obligation of having (or expressing) reactive attitudes is at least partially determined by moral wrongness. For example, John Rawls (1971) holds that when we explain our experience of having reactive attitudes, we necessarily have to invoke the concept of right and wrong.

The Reactive Attitude View

The Reactive Attitude View claims the normative status of having (or expressing) core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt) is conceptually prior to moral wrongness, but not vice versa (Mill, 1861; Rawls, 1971; Gibbard, 1990; Darwall, 2006; Skorupski, 2010a, 2010b).

If the Standard View adopts a top-down approach, then the Reactive Attitude View takes a bottom-up approach. This is because the former regards important moral categories as more basic than sundry empirical experiences, while the latter regards natural human practices as more basic than abstract moral categories. Proponents of the Reactive Attitude View believe that the notion of moral wrongness is very abstract and mysterious, so it is hard to figure out what one morally should or should not do. Yet, the notions of blame, resentment, etc. are concrete and familiar, since we naturally generate various and numerous emotions towards others and ourselves when we interact with people in our everyday life. Along this vein, they claim that we can make sense of moral wrongness by the normative status of our having (or expressing) reactive attitudes. For example, Skorupski (2010a, 2010b) defines moral wrongness in terms of the appropriateness of the sentiment of blame (namely, blameworthiness). In particular, he takes the following claim to be the definition of wrongness: “[i]t is morally wrong for x to a if and only if, were x to a from the beliefs that are warranted in x ’s epistemic state, then either x would be blameworthy for a -ing or extenuating circumstances would apply to x ’s a -ing.”⁸

The Independence View

The Independence View claims that neither moral wrongness nor the normative status of our having (or expressing) core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt) is conceptually prior to the other.

⁸ Skorupski, 2010b, p.292.

Two thoughts might motivate the Independence View. First, if one is skeptical of whether conceptual priority exists, and one take the Non-Naturalist Reactive Attitude Approach, then one is inclined to take the Independence View. Second, if one believes that we can grasp the concept of moral wrongness independently of the concept of the normative status of having or expressing core negative reactive attitudes, then one will reject the Reactive Attitude View. If one further believes that we can grasp the concept of such normative status independently of the concept of moral wrongness, then one will reject the Standard View. As a result, one will accept the Independence View.

6. Conclusion

My question is about the relationship between the concept of moral wrongness and the concepts of blame, resentment, indignation, and guilt. Reactive attitudes are the emotions an agent holds towards another agent or herself in response to what she believes to be the manner that the target agent exercised agency, whereas she regards the target agent as a fellow human being rather than a research object. They contrast with objective attitudes in that their holders believe the target agent to have exercised agency regarding the relevant action while holders of objective attitudes do not believe so. I will examine two approaches to the problem about the conceptual relationship between moral permissibility and core reactive attitudes: the Naturalist Reactive Attitude Approach and the Non-Naturalist Reactive Attitude Approach. For proponents of the Non-Naturalist Reactive Attitude Approach, they can take the Standard View, Reactive Attitude View, or the Independence View. With a basic understanding of these approaches and positions, I will

critically evaluate each. In the next chapter, I will reject the Naturalist Reactive Attitude Approach.

Ch. 2 The Naturalist Reactive Attitude Approach

In this chapter, I will argue against the Naturalist Reactive Attitude Approach. This approach semantically connects moral wrongness with the empirical disposition to hold/express core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt).⁹

This chapter will proceed as follows. First, I will present the Naturalist Reactive Attitude Approach. Second, I will motivate this approach. Third, I will show that this approach fails by arguing that there exist no semantic connection between moral wrongness and the empirical disposition to hold/express core negative reactive attitudes.

1. The Naturalist Reactive Attitude Approach

The core idea of the Naturalist Reactive Attitude Approach is that, given certain conditions, our empirical disposition to hold/express negative reactive attitudes is always connected with an agent's action being morally wrong. It makes sense when we think of paradigmatic cases. For example, we normally feel resentful towards an agent if we know that she murders our parent. Theorists endorsing the Naturalist Reactive Attitude Approach believe that, once we spell out the relevant conditions, most reacting people's negative reactive attitudes always connect with a target agent's action being morally wrong. Hence, these theorists all endorse the naturalistic reactive attitude thesis:

⁹ In principle, this approach can take up three different views: (1) moral wrongness is, at least in part, defined by such empirical disposition (2) such empirical disposition is, at least in part, defined by moral wrongness, and (3) neither moral wrongness nor the empirical disposition can be defined in terms of the other, although they are conceptually connected. Yet, since the empirical disposition to hold/express a reactive attitude is a psychological matter, (2) is implausible and I rule it out.

Naturalistic Reactive Attitude Thesis: It is morally wrong for an agent, S, to perform an action, A, if and only if almost all reacting people under certain conditions, C, would feel (or express) the blame-sentiment towards S regarding S's performing A, if they believed that S had no excuse.

What I mean by 'reacting people' are people who are capable of reacting to others and assessing the deontic status of actions. They need not be morally responsible agents, who are capable of bearing moral responsibilities. For example, a seriously depressed person may not be a morally responsible agent due to her abnormal mental state, but she can be a reacting person in that she has the capacity of assessing actions and reacting to others.

As I stated in Chapter 1, I focus on *feeling* or *having* reactive attitudes instead of *expressing* reactive attitudes. This is because the disposition to feel resentful, indignant, etc. is mostly backward-looking, dependent upon the target agent's past behaviors. In contrast, the disposition to express resentment, indignation, etc. is partly forward-looking, dependent upon the circumstances and impacts of the expression. For example, our disposition to *feel* resentful towards an agent who treated us unfairly mostly depends upon what she did to us, whereas our disposition to *express* resentment towards such agent depends not only on what she did, but also on the consequences of our expression, our and her well-being, the etiquette, etc. Moral wrongness is supposed to be a property of behaviors (in our context, the target agent's past behaviors), rather than a property of behaviors, the likely consequences of people's reactions, and etiquette, etc. Hence, it is more plausible to appeal to the backward-looking explanation, i.e., the disposition to *have*

reactive attitudes, than the forward-looking explanation.¹⁰ Hence, in the remainder of this chapter, I will only discuss the version of the Naturalistic Reactive Attitude Thesis that holds that the empirical disposition to *have* reactive attitudes is independent of the wrongness of an action without excuse.

David Hume (1739-40) is a prominent theorist of the Naturalist Reactive Attitude Approach. More precisely, Hume believes that moral wrongness is defined, at least in part, in terms of the empirical disposition to have reactive attitudes. To begin with, he believes that moral reactive attitudes or moral sentiments are very natural for us: “These sentiments are so rooted in our constitution and temper, that without entirely confounding the human mind by disease or madness, ‘tis impossible to extirpate and destroy them.”¹¹ Furthermore, he holds that we understand morality from feelings instead of reason: “Morality...is more properly felt than judg’d of.”¹² In particular, he maintains that we understand morality only through the reactive attitudes we have when we ignore our self-interests: “Tis only when a character is consider’d in general, without reference to our particular interest, that it causes such a feeling or sentiment, as denominates it morally good or evil.”¹³

¹⁰ Two clarifications are necessary. First, to say that moral wrongness is a property of an agent’s past action does not mean that the content of moral wrongness is solely in the past. This is because a consequentialist can agree that moral wrongness is property of a past action but maintain that its content is whether the action generates the best consequences, which stretches to the future.

Second, to say that moral wrongness is a property of an agent’s past action does not imply that moral wrongness cannot function in a forward-looking explanation. We might avoid performing an action because it is likely to be wrong, while wrongness remain as a property of the action only.

¹¹ Hume, 1739-40, 3.1.2.8.

¹² Hume, 1739-40, 3.1.2.1.

¹³ Hume, 1739-40, 3.1.2.4.

P. F. Strawson (1962) probably adopts the Naturalist Reactive Attitude Approach. To start with, he emphasizes that having reactive attitudes towards ourselves or each other is natural for human beings to the extent that we cannot suspend them altogether: “[W]e cannot, as we are, seriously envisage ourselves adopting a thoroughgoing objectivity of attitude to others as a result of theoretical conviction of the truth of determinism...”¹⁴ Second, he holds that reactive attitudes are “essential elements in the concepts of *moral* condemnation and *moral* responsibility.”¹⁵ Although his analysis of moral responsibility is not entirely clear, it seems that he finds the disposition of having reactive attitudes to be a part of the concept of moral responsibility. Given the conceptual connection between moral responsibility and moral wrongness, Strawson seems to accept that our empirical disposition to have reactive attitudes is connected to the concept of moral wrongness.

R. Jay Wallace (1994), on one construal, adopts a variant of the Naturalist Reactive Attitude Approach. When he addresses the problem of moral responsibility (i.e., what moral responsibility requires), he defines the practice of holding people morally responsible in terms of reactive attitudes. In particular, he argues that for us to hold an agent morally responsible is to hold her “to an expectation (or demand),” whereas to hold an agent to such expectations is “to be susceptible to the reactive emotions [i.e., resentment, indignation, and guilt] in the case that the expectation is breached, or to believe that the reactive emotions would be appropriate in that case.”¹⁶ Moreover, he

¹⁴ Strawson, 1962, p. 198.

¹⁵ Strawson, 1962, p. 206.

¹⁶ Wallace, 1994, p. 19, p. 51.

holds that moral wrongness is conceptually connected with this practice because “the notions of moral obligation, moral right, and moral wrong... form the nexus of concepts in which responsibility would seem to belong.”¹⁷ Accordingly, Wallace appears to hold that moral wrongness is connected either with our disposition to feel resentment, indignation, and guilt or connected with our belief that such attitudes are appropriate.

2. Motivations for the Naturalist Reactive Attitude Approach

Moral naturalists, who believe that moral facts and properties are natural, are inclined to take the Naturalist Reactive Attitude Approach is attractive for moral naturalists. These theorists hold that the objects of morality is as natural as the objects of science are. If so, then it is not surprising to construe moral phenomena as natural phenomena. More specifically, an action being wrong might just as well be our disposition to hold core negative reactive attitudes towards the agent.

A moral naturalist may be drawn to the Naturalist Reactive Attitude Approach because reactive attitudes are very natural to us when we react to people’s actions. In particular, we comprehend morality through our experiences. Having reactive attitudes is just the sort of experiences we have when we judge an action to be morally right or wrong. The notions of resentment, guilt, etc. are familiar and easy to grasp, and we are born to have them. It seems a truism when Hume says, “The mind of man is so formed by nature that, upon the appearance of certain characters, dispositions, and actions, it immediately feels the sentiment of approbation or blame; nor are there any emotions

¹⁷ Wallace, 1994, pp. 51-52.

more essential to its frame and constitution.”¹⁸ If having reactive attitudes is all we experience when we judge an action right or wrong, then the notions of right and wrong might turn out to be grounded by the notion of reactive attitude.

If morality is indeed naturalist, it may seem that wrongness is equivalent to having reactive attitudes, and hence that we have good reasons to accept the Naturalistic Reactive Attitude Thesis. However, I will argue that this thesis is false even if moral naturalism is true.

3. Problems of the Naturalist Reactive Attitude Approach

3.1 General Reasons against Moral Naturalism

Before I argue against the Naturalistic Reactive Attitude Thesis, I would like to comments on moral naturalism in general. First, it is quite controversial whether morality is naturalist, since a number of theorists have argued that moral rightness or wrongness is a non-natural property and has no naturalist definition. For instance, C. D. Broad (1928) claims that rightness cannot be defined in terms of natural terms (and hence not in terms of empirical reactive attitudes) in “Analysis of Some Ethical Concepts:”

I very much doubt whether ‘rightness’ can be defined. I am almost certain that it cannot be defined in non-ethical terms. [...] At any rate, I do not know, and cannot think of, any satisfactory definition [of rightness].¹⁹

¹⁸ Hume, 1751, 102.

¹⁹ Broad, 1928, p. 76.

W. D. Ross and H. A. Prichard agree with Broad that rightness has no naturalistic definition. Prichard (1937) claims that when we attribute a moral obligation to a person, what we attribute is *sui generis*, “incapable of having its nature expressed in terms of the nature of anything else,” i.e., anything non-moral.²⁰ Since moral obligation is conceptually connected to moral rightness/wrongness, we can infer that he does not deem the definition of moral rightness/wrongness to be naturalistic. In *Foundations of Ethics* (1939), Ross evaluates several theories attempting to define rightness naturalistically and concludes that they all fail. Recently, Russ Shafer-Landau, in *Moral Realism* (2003), defends the non-naturalist view of morality, claiming that moral facts are distinct in kind from natural facts.

As an argument against moral naturalism, G. E. Moore’s open question argument (1903) remains important. Moore points out that, for any naturalistic analysis of moral goodness, we can sensibly ask whether anything having the purported natural property is good. For example, for a proposal that moral goodness is being desired by a person, it is an open question whether something being so desired is morally good. This strategy applies to almost every moral term, since we can ask whether something is morally right/wrong even though it satisfies a relevant naturalist description.

It must be admitted that some theorists have levelled objections to the open question argument.²¹ For example, the reason why it is an open question to ask whether something being desired by a person is morally good might not be that moral goodness is not being so desired, but that, competent English language users do not have enough

²⁰ Prichard, 1937, p. 169.

²¹ Miller, 2003, pp. 15-18. Hurka, 2010.

knowledge about the notion of moral goodness. I don't regard the open question argument to be a decisive argument against moral naturalism. Nevertheless, it creates a burden for moral naturalists to justify their position.

Second, even for moral naturalists, many of them do not, at least explicitly, advocate moral wrongness in terms of the empirical disposition to hold core negative reactive attitudes.

For instance, Nicholas Sturgeon (1985) suggests that naturalism does not require a reductive naturalistic definition for a moral term. He thinks that it is compatible for a naturalist to hold that (i) moral facts are nothing but natural facts, and (ii) there is no reductive naturalistic definition for a moral expression.

Siding with Sturgeon, Boyd (1988) is optimistic in an adequate naturalistic definition of a moral term. Drawing on scientific realism (i.e., the view that scientific theories describe real phenomena that are largely independent of theorizing), he proposes that moral goodness could be a cluster of human goods that are homeostatically grouped. Along this line, he suggests that “[m]oral goodness is defined by this cluster of goods and the homeostatic mechanisms which unify them.”²² It is coherent for him to think that moral wrongness is also defined by a cluster of morally wrong actions.

Unlike Sturgeon and Boyd, Peter Railton (1986) offers a specific account of moral realism. According to him, the notion of moral rightness is “an idealization of the notion of social rationality by considering what would be rationally approved of were the interests of all potentially affected individuals counted equally under circumstances of

²² Boyd, 1988, p. 203.

full and vivid information.”²³ If approval is regarded to involve all positive reactive attitudes, then Railton basically defines moral rightness in terms of all the affected individuals’ disposition to hold positive reactive attitudes under full information. It follows that moral wrongness is basically the absence of such disposition, if there is no moral indeterminacy.

As we have seen, there is some reason to reject moral naturalism, and many moral naturalists do not openly advocate the Naturalist Reactive Attitude Approach. Although these considerations do not show the approach to fail, but they take away considerable motivations for the approach. In the next subsection, I will directly argue that the Naturalistic Reactive Attitude Thesis is false.

3.2 Problems of the Naturalistic Reactive Attitude Thesis

According to the Naturalistic Reactive Attitude Thesis, it is morally wrong for an agent, S, to perform an action, A, if and only if almost all reacting people under certain conditions, C, would feel (or express) the blame-sentiment towards S regarding S’s performing A, if they believed that S had no excuse for performing A. There is a constraint on C: the relevant description of C is not perfectly ideal. The constraint is necessary because theorists want to discover the empirical conditions for real people to satisfy the thesis, rather than confirm that the thesis holds under ideal conditions for hypothetical people. The former, as an empirical discovery, will prove the Naturalistic Equivalence Thesis and lend support to the Naturalist Reactive Attitude Approach, for it enables us to learn the purported connection between moral wrongness and natural facts. The latter, in contrast,

²³ Railton, 1986, p. 22.

is trivial to the truth of the thesis and make the Naturalist Reactive Attitude Approach collapse into the Non-Naturalist Reactive Attitude Approach. This is because, under perfectly ideal conditions, it is normative for all reacting people to have the blame-sentiment towards an agent who excusably acts wrongly, if these people believe that she has no excuse. In other words, under perfectly ideal conditions, what connects with moral wrongness is no longer the empirical disposition to have core negative reactive attitudes, but the normative status of having such attitudes.²⁴ In the following, I will show that no satisfying description of C will satisfy this constraint.

Let us explore of which C should consist. First, almost all the reacting people should have the empirical disposition to blame wrongdoers. If a reacting person can recognize a morally wrong action but feels indifferent or even favorable towards the wrongdoer, then her disposition to have negative reactive attitudes does not help indicate whether the relevant action is morally wrong. Hume correctly notices that we normally meet this requirement because we are generally sympathetic to others.²⁵ For example, when we watch a movie in a theater, we normally feel resentful towards those who deliberately mistreat the main character. However, in the meantime, we have the disposition to be more sympathetic to ourselves and people close to us, and less sympathetic to those who are distant from us. Thus, we have to make sure that such disposition does not get in the way when we react to people's actions, which makes the next requirement essential.

²⁴ This is a variant of the Euthyphro problem: do almost all the reacting people have the blame-sentiment towards the agent in question *because* they are under the perfectly ideal conditions, or are the conditions perfectly ideal *because* these people should have the blame-sentiment?

²⁵ Hume, 1751, 180.

Second, almost all the reacting people should be impartial in reacting to people's actions. That is, regardless of the relations between a reacting person, S_1 , and a target agent, S_2 , S_1 should almost always blame S_2 so long as S_2 acts wrongly and S_1 believes S_2 to have no excuse. Being aware of our natural biases, Hume requires us to neglect the differences between our interests and the interests of those who are remote from us when we react.²⁶ This requirement prevents us from unfairly favoring ourselves and our close ones and thus helps reduce the number of cases where reacting people have different reactive attitudes towards the same agent regarding the same action.

Third, almost all the reacting people should have the ability to recognize excusing conditions. For example, a reacting person should believe that an agent has excuses if that agent steals money under coercion.

Last, we have to impose some procedural constraints on the reacting people. For example, all the reacting people should have enough time to get the information about the action being reacted to, should not be influenced by any emotional distress, should not be in a rush to react to the relevant action, etc. These constraints serve to ensure that a reacting person does not feel any reactive attitude she does not have a disposition to have when she reacts in character.

When we put all these requirements together—the disposition to blame wrongdoers, being impartial, being capable of recognizing excusing conditions, and procedural constraints—it seems that we have successfully filled out the conditions that satisfy the Naturalistic Reactive Attitude Thesis, the conditions under which almost all reacting people would feel the blame-sentiment towards those who act wrongly when the

²⁶ 1739-40, 3.1.2.4; 1751, 186.

reacting people believe the target agents to have no excuse. However, I shall argue that the requirements we have so far are not sufficient to render the thesis true.

Think about the debate on abortion, for example. This debate has been going on for at least decades, but no unanimous judgments have been reached. Different stances are taken regarding when a fetus becomes a person in the moral sense. Some believe that a fetus is a person at conception because of their religious beliefs, and some believe that a fetus is a person when the fetus is sentient because of their deep beliefs about valuable lives, while others go as far as to regard a fetus as a person when the fetus is conscious. This fundamental disagreement about the point at which a fetus becomes a person causes people to react very differently towards a woman who has an abortion. If Humeans prefer the talk of feelings instead of beliefs, we can say that some are more sympathetic to the fetus while others are more sympathetic to the mother. Facing the divergent reactive attitudes, we can discount the reacting people who are not sympathetic to others, react rashly, or are partial, etc. However, among the remaining people, disagreement in judgment and reactive attitudes persist. The reason is that although people hold drastically different moral beliefs, there are no reliable ways to determine which beliefs are correct. For example, we have intuitions, coherence, etc. as rules of thumbs, but they cannot generate relatively indisputable results as scientific methods do. Hence, there is no guarantee that most of the reacting people after screening will have the same reactive attitude towards the same woman who has an abortion.

Perhaps morality is indeterminate in some cases, some might think. If so, then divergence in reacting people's reactive attitudes might correctly reflect the cases where the deontic status of an action is indeterminate. I agree that morality might be

indeterminate sometimes, but the number of cases where people react differently far outnumbers the possible cases where morality might be indeterminate. For example, if a fetus becomes a person at certain point, P, then abortion is morally indeterminate only when a fetus is approaching P, other things being equal. But we can reasonably believe that the cases where rational and impartial people react differently are much wider than cases where a fetus is approaching P. Hence, the appeal to moral indeterminacy does not help.

We can look at the problem in another way. It is a contingent matter of how many people hold particular religious beliefs. Suppose Christianity is the dominant religion, Christians typically are against same-sex marriage, and most of non-Christians are also against same-sex marriage. It follows that most of the reacting people (after screening) will feel the blame-sentiment towards same-sex couples, other things being equal. If the Naturalistic Reactive Attitude Thesis holds in this case, then it implies that same-sex marriage is morally wrong. However, as the number of Christians change and more non-Christians start to approve of same-sex marriage, most of the reacting people might not feel the blame-sentiment towards same-sex couples, other things being equal. If the Naturalistic Reactive Attitude Thesis still holds, it implies that same-sex marriage is morally right. Hence, it becomes a contingent matter whether same-sex marriage is morally right. But this implication is unacceptable.

Some might respond that morality is relative to the culture where the reacting people live. If they are right, then it is not surprising that whether same-sex marriage is morally right is contingent, for it is contingent which basic beliefs will be held by the culture where a reacting person lives. However, if morality is indeed relative, it cannot be

relative in such a narrow way, for it has many unacceptable implications. For example, it implies that reacting people in two different cultures, if they are impartial, cannot genuinely disagree about whether same-sex marriage is morally right. It also implies that the majority impartial moral judgments in a reacting culture are always correct, never making any progress. Because of these implications, I conclude that the move to relative morality does not work.

Then, can we stipulate that all reacting people have a set of correct fundamental moral beliefs? Doing so will probably solve the problem that reacting people have different reactive attitudes towards the same agent regarding the same action, for the same basic moral beliefs, the same disposition to be sympathetic to others, and the same impartial instance should generate the same reactive attitudes. However, this move violates the constraint of not making the reacting conditions perfectly ideal. The reason is that, since philosophers, who are supposed to be experts, have vastly different fundamental moral beliefs, it is hard for ordinary people to acquire correct fundamental moral beliefs. We could stipulate all reacting people to possess these beliefs, but that case is far too ideal, so ideal such that, along other requirements, it will make the Naturalistic Reactive Attitude Thesis trivially true. Thus, there is no way in which the reacting conditions in the thesis can be filled out in an informative way. Therefore, I conclude that the Naturalistic Reactive Attitude Thesis is false, which entails that the Naturalist Reactive Attitude Approach is false.

4. Conclusion

In this chapter, I have rejected the Naturalist Reactive Attitude Approach to the conceptual relationship between moral wrongness and reactive attitudes. According to this approach, an action's being morally wrong must be connected with the presence of almost all reacting people's core negative reactive attitudes under specified conditions. I have argued that the idea above is false because the relevant conditions cannot be specified in a substantive way. The failure of the Naturalist Reactive Attitude Approach gives us a reason to adopt the Non-Naturalist Reactive Attitude Approach, which connects moral wrongness and the normative status of hold negative reactive attitudes. This is the approach I will defend in the next chapter.

Ch. 3 The Non-Naturalist Reactive Attitude Approach

In this chapter, I will defend the Non-Naturalist Reactive Attitude Approach. This approach semantically connects moral wrongness with the normative status (appropriateness, moral permissibility, etc.) of having/expressing core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt). Theorists of this approach can take three different views: (1) the Standard View, moral wrongness is conceptually prior to the normative status of having/expressing reactive attitudes, but not vice versa (2) the Reactive Attitude View, which holds that by such normative status is conceptually prior to moral wrongness, but not vice versa, and (3) the Independence View, neither moral wrongness nor such normative status is conceptually prior to the other. I maintain that the Standard View is correct. I will first argue for the Non-Naturalist Reactive Attitude Approach in this chapter, then argue against the Reactive Attitude View and the Independence View in the next chapter. After that, I will proceed to argue for the Standard View.

This chapter will proceed as follows. First, I will present the Non-Naturalist Reactive Attitude Approach. Second, I will motivate this approach. Third, I will respond to critics of this approach.

1. The Non-Naturalist Reactive Attitude Approach

According to the Non-Naturalist Reactive Attitude Approach, moral wrongness is semantically connected to the appropriateness or permissibility of holding/expressing a reactive attitude. Thus, presumably we can understand moral wrongness in terms of such appropriateness, or we can understand such appropriateness in terms of moral wrongness.

Theorists of the Non-Naturalist Reactive Attitude Approach all accept the following thesis:

Non-Naturalistic Reactive Attitude Thesis: It is morally wrong for an agent, S, to perform an action, A, if and only if were S to have no excuse then it would be appropriate (permissible, or obligatory) for any reacting person to feel (or express) the blame-sentiment towards S regarding S's performing A.

The Non-Naturalist Reactive Attitude Approach is *not* non-naturalist in the sense that it regards moral properties to be *sui generis* non-natural properties. Rather, it is non-naturalist precisely in the sense that it takes moral wrongness and reactive attitudes to be connected in a non-natural way. Hence, those endorsing this approach need not be moral non-naturalists. The theorists adopting the Non-Naturalist Reactive Attitude Approach include J. S. Mill (1861), John Rawls (1971), Peter Railton (1986), Allan Gibbard (1990), Stephen Darwall (2006), John Skorupski (2010a, 2010b), and arguably include W. D. Ross (1930, 1939) and Immanuel Kant (1785, 1788). This approach also squares well with some other theorists' accounts, although they do not explicitly advocate it (e.g., Nicholas Sturgeon 1985; Richard Boyd, 1988; David Brink 1989; Russ Shafer-Landau, 2003; David Enoch, 2011).

Some of these theorists hold the Reactive Attitude View, the view that the normative status of having/expressing core negative reactive attitudes is conceptually prior to moral wrongness, but not vice versa. For example, Mill accounts for moral

wrongness in terms of the moral obligation to express blame and punishment. He famously says,

For the truth is, that the idea of penal sanction, which is the essence of law, enters not only into the conception of injustice, but into that of any kind of wrong. We do not call anything wrong unless we mean to imply that a person ought to be punished in some way or other for doing it—if not by law, by the opinion of his fellow creatures; if not by opinion, by the reproaches of his own conscience.²⁷

What Mill has in mind is unclear, but the quote suggests that he understands moral wrongness in terms of the conclusive obligation to express blame or punish the agent, since he talks about punishment and what we ought to do to the target agent.²⁸

More recently, Darwall proposes that moral wrongness is analyzable by blameworthiness, the appropriateness of having the blame-sentiment: “What is morally wrong is what is blameworthy—that is, what is warrantably blamed, if the action is done without adequate excuse.”²⁹ ‘Warrant’ here roughly refers to a fitting relationship between the action and the blame-sentiment. More precisely, Darwall holds that a morally wrong action just is the action regarding which the agent is a fitting object of the blame-sentiment for every member of the moral community, if the agent lacks excuses.

²⁷ Mill, 2001, pp. 48-49.

²⁸ Brink, D, 2008.

²⁹ Darwall, 2010, p. 143.

In contrast, some theorists taking the Non-Naturalist Reactive Attitude Approach accept the Standard View, the view that the normative status of having/expressing reactive attitudes is conceptually prior to defined by moral wrongness, but not vice versa.

For instance, Rawls states,

[I]t is a necessary feature of moral feelings, and part of what distinguishes them from the natural attitudes, that the person's explanation of his experience invokes a *moral* concept and its associated principles. His account of his feeling makes reference to an acknowledged *right* or *wrong*.³⁰

Since moral feelings are reactive attitudes in Rawls's view, he takes reactive attitudes to presuppose rightness/wrongness. It seems to follow that holding/expressing an attitude can be appropriate or inappropriate because there is an objective fact about whether the target agent inexcusably performed a wrong action.

For the reasons discussed in the previous chapter, I will focus on the version of the Non-Naturalistic Reactive Attitude Thesis that holds the appropriateness, (permissibility, or obligation) of *having* reactive attitudes is connected to the wrongness of an action. Only the *appropriateness* of having reactive attitudes will be discussed, because my interest is in exploring the plausibility of a connection between moral wrongness and the normative status of having reactive attitudes. Since obligation is simply permissibility plus wrongness of the negation, and permissibility is simply the negation of wrongness, such accounts would be circular as accounts of wrongness. So, I

³⁰ Rawls, 1971, p. 481, my emphasis.

will merely discuss the version of the Non-Naturalist Reactive Attitude Approach that holds that *appropriateness* of *having* reactive attitudes is connected to the wrongness of an action.

We have seen that the Non-Naturalist Reactive Attitude Approach is quite popular. In the next section, I will explain why so many theorists are attracted by this approach.

2. Motivations for the Non-Naturalist Reactive Attitude Approach

The Non-Naturalist Reactive Attitude Approach is attractive at least for two reasons. First, we have seen that Naturalist Reactive Attitude Approach fails partly because it has extreme difficulty in specifying the conditions that always connect the disposition to have the blame-sentiment with moral wrongness. Perhaps the problem with this account is not that there is no conceptual relationship between moral wrongness and reactive attitudes, but that the naturalist reactive attitude approach is not the right sort of approach. If so, then the Non-Naturalist Reactive Attitude Approach might be the right approach.

Second, and this is for people proposing the Reactive Attitude View only, it is promising to explicate a relatively mysterious notion (namely, the notion of moral wrongness) in terms of a relatively familiar notions (namely, the notion of blame, resentment, indignation, and guilt, etc.). We are quite familiar with blame, resentment, indignation, and guilt, etc. because we commonly see people hold or express these attitudes. In contrast, we do not “see” moral wrongness in actions, but we believe some actions to be morally wrong. Since morally wrong actions are often correlated with reactive attitudes, it is plausible that some essential connection can be made between moral wrongness and the normative status of holding (or expressing) reactive attitudes.

I side with the Non-Naturalist Reactive Attitude Approach. I would like to directly argue for the Non-Naturalistic Reactive Attitude Thesis, but I can find no non-circular argument to do so. This is because I think that the appropriateness of holding a negative reactive attitude, semantically speaking, is just the wrongness of the relevant action and the absence of excuse. Hence, in the next section, I will indirectly argue for the thesis by showing that opposing views of the thesis are wrong.

3. Replies to Critics

A number of views are opposed to the Non-Naturalistic Reactive Attitude View. Some opposing views might say that the appropriateness of holding a reactive attitude is connected with the reacting person's beliefs, for instance. For simplicity, I will talk about blameworthiness in general, instead of the appropriateness of holding a negative reactive attitude.

First, an opposing view is ruled out because of the sense of appropriateness I am concerned about. Recall that I distinguish between objective and subjective appropriateness, where objective appropriateness is the appropriateness regarding the real situation and subjective appropriateness is the one regarding the situation believed by the reacting person.³¹ Thus, if it is *subjectively* appropriate for Reactor to harbor resentment towards Agent due to Reactor's misunderstanding of Agent's behavior, it might not be *objectively* appropriate for Reactor to feel so.

Now let us examine the opposing views concerning objective appropriateness. If the thesis is false, then the determining factors of objective appropriateness are not, or at

³¹ Chapter 1, Section 2.

least not exclusively, deontic status and the presence of excuse. To me, the prominent candidates bearing upon objective appropriateness are the target agent's beliefs, the target agent's intentions, the reacting person's evidence, the consequences of the relevant action, and all available actions.

Some might hold that whether an agent is blameworthy mainly depends upon the agent's mental state, either her beliefs or her intentions. Michael Zimmerman, for example, holds that an agent's being blameworthy depends requires her belief that she is doing wrong.³² Thus, he believes that Huckleberry Finn is blameworthy for having freed his slave friend Jim because Huckleberry thought of himself to be doing the wrong thing. Zimmerman reminds us that "the agent is indeed doing wrong *from his perspective* even if not in fact..."³³

I agree that Huck is doing wrong from his perspective, and there is something criticizable if an agent acts despite her belief that she is doing wrong. However, I do not think that any reacting person is entitled to blame Huck. Because Zimmerman's account of blameworthiness requires a target agent to have the belief that she is acting wrongly, he will regard an agent not to be blameworthy if she acts wrongly with the belief that she is acting rightly, other things being equal. This is clearly false. How can a driver not be blameworthy if she voluntarily gets drunk, drives home, while believing herself to be acting rightly the whole time? She acts wrongly, and her beliefs about her acts do not

³² He seems to use 'blameworthy' and 'culpable' interchangeably. Zimmer's full account of blameworthiness is this: "Every chain of [blameworthiness] is such that at its origin lies an item of behavior for which the agent is directly culpable and which the agent believed, at the time at which the behavior occurred, to be overall morally wrong." Zimmerman, 2006, p. 176.

³³ Zimmerman, 2006, p. 201.

constitute any excuse for her acting wrongly, so she is blameworthy. This serves as a counterexample to the view that the belief-based view.

Some base blameworthiness upon intentions. The thought is that an agent is not blameworthy if she acts wrongly but acts with good intentions. For instance, assuming consequentialism, if Helper intends to help an old lady cross the street and ends up scaring the lady away because of his aggressive appearance, then it seems that it is inappropriate to blame Helper even if he acts wrongly. Note that this case is supposed to constitute a counterexample to the Non-Naturalistic Reactive Attitude Thesis only if consequentialism is true. If consequentialism is false, then Helper's action may be right, so the case will not falsify the thesis.

I agree that we assess an agent's mental states. For example, we draw a distinction between attempted murder and accidental murder because an agent commits attempted murder has a guilty mind while the agent commits accidental murder does not. Yet, even when we draw this distinction, we first make sure that the agent did wrong. In any case, we do not punish anyone with a guilty mind but no guilty acts. At best, intentions might serve as partial excuses for doing wrong.

In cases like the Helper, the relevant excuse is the agent's ability to foresee the likely consequences, not his intentions. If Helper could not have foreseen that his look will scare away the lady he intends to help, then we can say that he has an excuse for acting wrongly, which renders him not blameworthy. But if he has good reasons to believe that his look is scary for elders, then even though he has good intentions, he still has no excuses for acting wrongly. Thus, with this twist, he appears to be blameworthy.

In other words, the fact that has good intentions simply cannot exculpate him. If I am right, then the intention-based view does not work as an account of blameworthiness.

A third opposing view we will consider holds that whether an agent is blameworthy depends upon the evidence available to the target agent. That is, a target agent is blameworthy if and only if the evidence to her suggests that she is acting wrongly. For example, Student steals a book from the library with no excuse, but her evidence suggests that she is acting rightly (e.g., it is a valuable book that will be destroyed soon). In this case, it seems that it is not appropriate for us to blame her regarding the missing book.

This evidence-based view works if the target agent's evidence is all the possible evidence she can reasonably have. In the Student case, for instance, Student's evidence might function as excuses and render her not blameworthy. However, when a target agent's evidence suggests that she is acting rightly, there could be a few pieces of evidence she can easily get and make her seriously doubt whether she is acting rightly. Accordingly, she does not appear to be innocent, and hence she is blameworthy. Hence, the evidence-based view is too narrow.

The fourth view holds that an agent is not blameworthy if she acts wrongly but the action produces good consequences. (Such actions will not falsify the Non-Naturalistic Reactive Attitude Thesis for consequentialists because they will regard such actions as morally right.) For example, suppose Husband acts wrongly in lying to Wife without any excuse, but the lie accidentally and unexpectedly makes Wife feel happy and she will never find out about the lie. Some might argue that Husband is not blameworthy,

and may even be praiseworthy, for her action due to the good consequences he brought about.

Alternatively, an agent may act rightly but the action produces bad consequences. Some might argue that the agent should be blamed for performing the action although the action is right. (Again, such cases are compatible with the thesis for consequentialists because they will regard the relevant action morally wrong.) For instance, Daughter truthfully and rightly tells her mother that her father has an affair, which results in her parents' divorce, which is bad for each. Some might hold that Daughter should be blamed for telling the truth.

I maintain that good consequences do not ensure lack of blameworthiness. The blameworthiness in question is based on the relevant action's deontic status. In the Husband case, if he acts wrongly, then he is blameworthy. And in the Daughter case, if she acts rightly, then she is not blameworthy.

Nevertheless, some will question why we still feel that consequences matter for our reactions to the agent. In any case, it seems that Husband's making Wife happy is relevant for how we should treat him, and Daughter's causing her parents' divorce is relevant for how we should treat her.

To this question, I claim that an action's consequences affect our reactions to the agent because we care about the changes in our life, which changes include the changes an agent has brought about. We naturally have attitudes towards the changes in our life, and thus we have attitudes towards who has brought about the changes, be it us, our friends, or strangers. Moreover, what change an agent has brought about, to some extent, indicates the changes she may bring about in the future. So, the consequences of an

agent's action matter to us because they are part of the changes of our life (and the life of the ones we care). However, just because consequences matter to us, it does not mean that the agent at stake is praiseworthy or blameworthy regarding the action.

However, just because we have attitudes towards an agent regarding her action's consequences, it does not mean that we need to blend such attitudes into our reactive attitudes towards an agent regarding her action's deontic status. These attitudes are distinct and we can separate them in principle. We can maintain that Husband is as blameworthy as if he lied without making Victim happy, although we are more satisfied with the Wife being happy than she being unhappy. Likewise, Daughter is as not blameworthy as if she told the truth without causing her parents to divorce, although we are more satisfied with her parents not divorcing than their divorcing.

If consequences do not matter in terms of objective appropriateness of reactive attitudes regarding deontic status, maybe the available actions do. For instance, if an agent acts rightly but other actions available to the agent produce much better consequences, then it seems that the agent is to blame in some sense. Suppose it is permissible not to donate any money to the International Red Cross. Even so, it seems that we can reproach Billionaire for his failure to donate, just because he is one of the ones with the best capacity to donate but still fails to do so.

In such cases, I claim, our reproach is towards the kind of person the agent exhibits through his action instead of deontic status of the action. What makes us think that Billionaire deserves reproach is not that the failure to donate money is purely wrong. Rather, it is that he, *as a billionaire*, is not willing to spare even a cent to the people in need. Put another way, he is a selfish person, and we think that selfishness is a morally

flawed character. In general, we tend to reproach people with morally flawed characters if they are morally responsible for developing their characters. Hence, this case fails to show the Non-Naturalistic Reactive Attitude Thesis to be false as well.

The most seemingly problematic cases for the thesis come from prohibition dilemmas. In a prohibition dilemma, all available actions are wrong. Some might argue that when an agent is in a prohibition dilemma, although she acts wrongly, it is not appropriate to hold the blame-sentiment towards her regarding her action. After all, “she has no choice,” “she has to do wrong,” etc.

I agree that it is sometimes inappropriate to blame an agent regarding her action in a prohibition dilemma, but I argue that if blame is inappropriate, then an excuse must exist. A familiar example is Sophie’s choice. At a Nazi concentration camp, Sophie is forced to choose which one of her two children will live and which one will be killed. If she chooses neither, then both will be killed. In this case, regardless of which child Sophie saves, she makes the other get killed. As a result, she acts wrongly regardless which child she chooses to save, for she is morally accountable for an innocent person’s death. However, she is not blameworthy because she has an excuse, namely, she is under severe coercion from the Nazi’s threat to kill both of her children, if she does not pick one. In general, one has an excuse if one is under coercion. Hence, she acts wrongly and has an excuse, and it is not appropriate to blame her, which fails to show the Non-Naturalistic Reactive Attitude Thesis to be false. Generally, in a prohibition dilemma, the agent in question has to act wrongly, so there is no justification for her. Yet she may have an excuse, for she may be under coercion, lacks or is not supposed to have the relevant knowledge, or is not competent, etc.

We need to note that, in prohibition dilemmas, excuses only exist when one is forced into the dilemma. For example, if Sophie chooses to swim with her two children, and her children are drowning, then she placed herself into a prohibition dilemma: either to save the older child and let the younger one get killed, or to save the younger child and let the older one get killed. In this case, Sophie has no excuse because she is responsible for the tragedy of her prohibition dilemma.

Such cases do not show the Non-Naturalistic Reactive Attitude Thesis false, since the thesis claims that if an agent inexcusably acts wrongly, then she is blameworthy for her action. Since Sophie voluntarily gets herself into a prohibition dilemma, she acts wrongly and she has no excuse. In addition, she is blameworthy for failing to save a child. Hence, this case is consistent with the thesis.

Thus far, I have argued that all standard objections against Non-Naturalistic Reactive Attitude Thesis fail. Since I regard these objections to be the strongest one, I conclude that the thesis is true.

4. Conclusion

In this chapter, I have defended the Non-Naturalist Reactive Attitude Approach to the conceptual relationship between moral wrongness and reactive attitudes. According to this approach, an action's being morally wrong is semantically connected with the presence of any reacting person's appropriate core negative reactive attitudes when she believes the target agent to lack excuses. I have argued that the idea above is true by arguing against the main opposing views.

In the next chapter, I will argue that the Reactive Attitude View is false. That is, I will argue that the normative status of holding/expressing core negative reactive attitudes is not prior to moral wrongness.

Ch. 4 The Reactive Attitude View

In this chapter, I will argue against the Reactive Attitude View. The Reactive Attitude View holds that moral wrongness is defined, at least partly, by the normative status of our having (or expressing) core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt), but not vice-versa.

1. The Reactive Attitude View

The Reactive Attitude View first appears in John Stuart Mill's *Utilitarianism* (1861). It is recently revived by Allan Gibbard (1990, 2003, 2008a, 2008b), Stephen Darwall (2006, 2010), and John Skorupski (2010a, 2010b). All these theorists maintain that moral wrongness is defined, at least partly, by the normative status (e.g., appropriateness, permissibility, *pro tanto* or conclusive moral obligation) of our having or expressing reactive attitudes.³⁴ Precisely, the Reactive Attitude View has this form:

Reactive Attitude View: It is morally wrong for an agent, S, to perform an action, A, if and only if, and in virtue of the fact that, were S to have no excuse then it would be appropriate (permissible, or obligatory) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A.

³⁴ Gibbard's account is a bit different from Darwall's and Skorupski's since he is interested in the subjective sense of moral wrongness (roughly, moral wrongness given the agent's epistemically justified beliefs). More on this later.

In the following, I will introduce different theorist's accounts by the normative status they appeal to: moral obligation (either *pro tanto* or conclusive), moral permissibility, or appropriateness.

1.1 Appeal to Moral Obligation or Moral Permissibility

Mill accounts for moral wrongness in terms of the moral obligation to express blame and punishment. He famously says,

For the truth is, that the idea of penal sanction, which is the essence of law, enters not only into the conception of injustice, but into that of any kind of wrong. We do not call anything wrong unless we mean to imply that a person ought to be punished in some way or other for doing it—if not by law, by the opinion of his fellow creatures; if not by opinion, by the reproaches of his own conscience.³⁵

There is controversy over what Mill means, but he seems to understand moral wrongness in terms of the conclusive obligation to express blame or punish the agent, since he talks about punishment and what we ought to do to the target agent.³⁶ Note that what he has in mind is moral wrongness relative to the reacting people's epistemically justified beliefs, since he does not consider excuse to be relevant at all. We can call this sense of moral wrongness 'subjective moral wrongness.' It is distinct from the sense of moral wrongness that is relative to all the facts, which we can call 'objective moral wrongness.' The sense

³⁵ Mill, 2001, pp. 48-49.

³⁶ Brink, D., 2008.

of moral wrongness we usually talk about is objective. In what follows, I refer to objective moral wrongness when I don't specify the sense of moral wrongness.

Mill's account is problematic in several respects. First, it appeals to expressing an attitude instead of having an attitude. As discussed in Chapter 2, it is more plausible to appeal to having an attitude than expressing an attitude. Roughly, this is because the disposition to have an attitude is mostly backward-looking, which matches moral wrongness, presumably a property of past actions. In contrast, the disposition to express an attitude is partly forward-looking, which does not match the backward-looking character of moral wrongness.³⁷

Second, this account appears circular because conclusive obligation (to express blame or punish) is used to explain subjective moral wrongness. Yet a conclusive obligation to do something conceptually entails the objective moral wrongness of not doing it, and objective moral wrongness amounts to subjective moral wrongness when the agent's epistemically justified beliefs are all true. However, this account is not circular if the utilitarian principle is introduced as a definition of conclusive obligation, where the principle is that one has a conclusive obligation to perform an action if and only if that action maximizes total utility. When we combine Mill's notion of subjective moral wrongness and the principle of utility, we can derive a non-circular notion of subjective moral wrongness, i.e., it is subjectively morally wrong for an agent, S, to perform an action, A, if and only if, and in virtue of the fact that, it is such that others' blaming or punishing S for performing A maximizes total utility. This account can be converted to the following account of objective moral wrongness: it is (objectively)

³⁷ Detailed explanations are in Chapter 2, Section 1.

morally wrong for an agent, S, to perform an action, A, if and only if, and in virtue of the fact that, it is such that others' blaming or punishing S for performing A maximizes total utility and others' epistemically justified beliefs are all true.

Rendered this way, Mill's notion of wrongness is not circular anymore, but his account ceases to be the Reactive Attitude View, for the obligation to blame is not doing any real work. Furthermore, this account generates problematic cases. In particular, a utility-maximizing action is morally right according to utilitarianism, but on Mill's account, it is morally wrong when others' blaming or punishing the agent for performing it maximizes total utility and others' epistemically justified beliefs are all true. An action cannot be both morally right and wrong. Hence, it does not seem to help when we insert the utilitarian principle to Mill's original account.

Admittedly, the problem of problematic cases will disappear if we apply the utilitarian principle to the action of blaming or punishing only. Accordingly, Mill's view has two levels. At the first level, the moral wrongness of the action of blaming or punishing is defined by whether the action maximizes total utility given the reacting people's epistemically justified beliefs are all true. At the second level, the moral wrongness of any action *other than* blaming and punishing is defined by whether it is morally wrong to blame or punish the agent for performing the action. However, this approach does not solve the problem of abandoning the Reactive Attitude View, since the obligation to blame or punish does not play any real explanatory role for the notion of wrongness any more.

In a nutshell, accounting for wrongness in terms of moral obligation is either circular or internally problematic. It is circular, since wrongness is indeed connected with

obligation. When the circularity is eliminated by offering an independent definition of obligation, the account ceases to be a Reactive Attitude View.

Like obligation, permissibility is conceptually connected with wrongness, so the problem of accounting for wrongness in terms of *obligation* simply transfers to accounting for wrongness in terms of *permissibility*. Accordingly, accounting for wrongness in terms of permissibility does not work for any Reactive Attitude View. Moreover, it does not matter whether wrongness is explained in terms of *pro tanto* or conclusive obligation/permissibility, for either is conceptually connected with wrongness. Thus, for the advocates of the Reactive Attitude View, it is more plausible for them to resort to the *appropriateness* of having reactive attitudes than the *permissibility* or *obligation* of having reactive attitudes. Let us call the version that resort to appropriateness ‘Fitting Attitude Theories of Wrongness.’

1.2 Appeal to Appropriateness

All Fitting Attitude Theories of Wrongness reduce moral wrongness to the appropriateness of having reactive attitudes. ‘Appropriateness’ has no deontic conceptual implications. Some use ‘proper,’ ‘fitting,’ ‘correct,’ etc. for the same purpose. I will use ‘appropriate’ and ‘fitting’ interchangeably.

In what follows, I will introduce Darwall’s, Skorupski’s, and Gibbard’s Fitting Attitude Theories of Wrongness. Darwall’s and Skorupski’s accounts reduce wrongness to a generic notion of blameworthiness, while Gibbard reduces wrongness to the appropriateness of having specific reactive attitudes (resentment and guilt in his

unrefined version, and anger and guilt in his refined version). That is, Darwall and Skorupski invoke blame only, while Gibbard invokes specific negative reactive attitudes.

Darwall reduces moral wrongness to blameworthiness (the appropriateness of having the blame-sentiment): “What is morally wrong is what is blameworthy—that is, what is warrantably blamed, if the action is done without adequate excuse.”³⁸ ‘Warrant’ here roughly refers to a fitting relationship between the action and the blame-sentiment. Put more fully, in Darwall’s view, a morally wrong action just is the action regarding which the agent is a fitting object of blame for every member of the moral community, provided that the agent lacks excuses.

Skorupski agrees with Darwall that moral wrongness is reduced to blameworthiness. In particular, he claims that wrongness is non-circularly defined in terms of blameworthiness. He takes the following claim to be the definition of wrongness:

- (1) It is morally wrong for x to a if and only if, were x to a from the beliefs that are warranted in x ’s epistemic state then either x would be blameworthy for a -ing or extenuating circumstances would apply to x ’s a -ing.³⁹

Extenuating circumstances, he explains, “remove or diminish blame that would otherwise have been appropriate.”⁴⁰ That is, extenuating circumstances are cases of excuses. That a

³⁸ Darwall, 2010, p. 143.

³⁹ Skorupski, 2010b, p. 292. He explains (1) in terms of (1^a): It is morally wrong for x to a =_{Df} were x to a from the beliefs that are warranted in x ’s epistemic state, and without any extenuating circumstances, there would be sufficient reason to δ x for a -ing, p. 295. Whereas δ refers to “the sentiment that is (i) the appropriate response to certain representative examples of belief and action; and that (ii) prompts a characteristic-type of action, namely, withdrawal of recognition,” p. 295. I will discuss Skorupski’s account more fully in Section II.2.

belief is warranted in an agent's epistemic state, for Skorupski, means that the agent is epistemically justified in having the belief.⁴¹ Because Skorupski claims that "[m]oral wrongness turns neither on the agent's actual beliefs, nor on the facts, but on the beliefs warranted in the agent's epistemic state,"⁴² he is discussing subjective moral wrongness. We can convert his account into an account of objective moral wrongness: It is morally wrong for x to a if and only if, were x to a , either x would be blameworthy for a -ing or extenuating circumstances would apply to x 's a -ing.

Darwall's and Skorupski's accounts differ in their explanations of blameworthiness. Darwall takes blameworthiness to imply "*demands* that are implicitly *addressed* to their attitudes' objects second-personally, if only in imagination."⁴³ What is essential to Darwall is that there exists certain relation between the reacting person and the target agent, which relation he dubbed "second-personal relation." When two people are in this relation, the reacting person could demand the target agent to perform an action, which demand gives the target agent a normative reason to act accordingly.⁴⁴ In contrast, Skorupski explains blameworthiness as the appropriateness of blame, whereas

⁴⁰ Skorupski, 2010b, p. 292.

⁴¹ Skorupski, 2010b, Ch. 5. He is an internalist about epistemic justification, that is, he thinks that whether an agent's belief is justified supervenes on the mental state of the agent. This position does not concern us here.

⁴² Skorupski, 2010b, p. 297.

⁴³ Darwall, 2010, p. 142, emphasis original.

⁴⁴ More on this in Darwall, 2006. Strictly speaking, the target agent and the agent holding the attitude could be the same person.

he defines blame by a purportedly characteristic action it disposes, i.e., “withdrawal of recognition.”⁴⁵

Gibbard is a non-cognitivist because he considers moral judgments to express commitment to plans, not propositions. The problem I am interested in does not bear on the cognitivism/non-cognitivism debate, for all the possible positions can adopt either cognitivism or non-cognitivism. Hence, I will bypass Gibbard’s non-cognitive stance when I discuss his account. Gibbard agrees with Darwall and Skorupski that moral wrongness is reduced to the generic notion of blameworthiness, but Gibbard does not stop the reduction there. Rather, he picks the warrant of resentment and guilt as the ground of moral wrongness.⁴⁶ Furthermore, he explicitly distinguishes subjective and objective moral wrongness and analyzes subjective wrongness by the warrant of resentment and guilt. In his view, an action is objectively wrong “if it is wrong in light of all the facts, knowable and unknowable,” whereas it is subjectively wrong “if it is wrong in light of what the agent had good reason to believe.”⁴⁷

Like Skorupski, Gibbard is interested in analyzing subjective wrongness only. Roughly, his analysis of subjective wrongness is this: an action is subjectively morally wrong if and only if “feelings of resentment or outrage over it are warranted on the part of impartial onlookers and feelings of guilt over it are warranted on the part of the person

⁴⁵ Skorupski, 2010b, p. 294.

⁴⁶ Gibbard talks about the rationality of guilt and resentment in Gibbard, 1990. He later switches to the talk of warrant (Gibbard, 1992b, 2006c, 2008a).

⁴⁷ Gibbard, 1990, p. 42.

who does it.”⁴⁸ We can convert this analysis into an account of objective wrongness in this way: an action is objectively morally wrong if and only if, were both the agent and the reacting people fully informed, both the agent’s feeling guilty for having done it is warranted and the reacting people’s feeling resentful towards her for having done it is warranted.⁴⁹ For Gibbard, a feeling is warranted just in case we plan to feel it as an answer to the normative question what to feel.⁵⁰ Alternatively, he suggests understanding warrant as reasons: “[t]he *rational* [warranted] thing to do [or to feel], we can say, is the thing to conclude by reasoning to do [or to feel], *if* you are to reason on the matter at all,” while “[t]o conclude how to reason is to plan for how to reason.”⁵¹

So far, I have rejected, versions of the Reactive Attitude View that appeal to moral *permissibility/obligation* to have (or express) reactive attitudes, and to the

⁴⁸ Gibbard, 2006c, p. 196.

His full account is this: [an] act is *wrong* if and only if it violates standards for ruling out actions, such that if an agent in a normal frame of mind violated those standards because he was not substantially motivated to conform to them, he would be to blame. To say that he would *be to blame* is to say that it would be rational for him to feel guilty and for others to resent him,” 1990, p. 45. The difference between the full account and the rough account is that the former requires certain psychological capacity from the agent while the latter does not have that requirement. For the purpose of this dissertation, we can ignore the psychological requirement and use Gibbard’s rough account for simplicity.

Gibbard later talks about moral right/wrong in terms of warranted feelings. Specifically, he said that to think an act morally wrong roughly is “to think that the act warrants resentment on the part of others and guilt on the part of the person who did it.” Gibbard, 2008, p. 16. See also Gibbard, 1992b. This formulation is semantically equivalent to his earlier view in Gibbard, 1990. I will mainly talk about his view in Gibbard (1990) for simplicity.

⁴⁹ Gibbard (2005) proposes a definition of the objective ought in terms of the subject ought: “what one ought_{ob} to do is what one would ought to do if one had full information. It’s what it would be the case that one ought to do, in the subjective sense, with full information.” Gibbard, 2005, pp. 345-346. He has a refined version of the definition, but the rough definition suffices my need for my purpose.

⁵⁰ Gibbard, 2006c, 2008a. In earlier works, he talks about warrant or rationality as accepting norms, 1990, 1992b.

⁵¹ Gibbard, 2003, p. 153, emphasis original.

appropriateness of *expressing* reactive attitudes. What are left are the ones appealing to the *appropriateness* of *having* reactive attitudes, namely, Fitting Attitude Theories of Wrongness. I have introduced three such accounts: Darwall's, Skorupski's, and Gibbard's. In the next section, I will motivate the Reactive Attitude View in general.

2. Motivations for The Reactive Attitude View

The Reactive Attitude View is attractive for mainly two reasons. First, it seems plausible that a relatively mysterious notion (namely, the notion of moral wrongness) is partly grounded by relatively familiar notions (namely, the notion of blame, resentment, indignation, and guilt, etc.). We are quite familiar with blame, resentment, indignation, and guilt, etc. because we commonly see people hold or express these attitudes. Having reactive attitudes is part of our daily life. In contrast, we cannot "perceive" moral wrongness from an action in the way we perceive an action to exist. It is more plausible to say that we believe, intuitively or not, some actions to be morally wrong. Since morally wrong actions are often correlated with reactive attitudes, it seems plausible that a conceptual grounding relation can be made between moral wrongness and the normative status of holding (or expressing) reactive attitudes.

Second, fitting attitude theories of value are popular nowadays. Seeing the appeal of such theories, we might conjecture that moral rightness/wrongness, like value, can be analyzed by fitting attitudes. According to fitting attitude theories of value, an apple is good because it is a fitting, apt, appropriate, or correct object of valuing for us. According to the theory, the value of the apple is reduced to the reasons that make the apple a fitting object of valuing for us. When we apply this reasoning to deontic concepts, we may find

it appealing to explain moral wrongness in terms of the reasons that make an agent a fitting object of blame, resentment, indignation, and guilt, etc. for reacting people. In fact, Gibbard and Skorupski subscribe to both fitting attitude theories of value and fitting attitude theories of wrongness.

Although the Reactive Attitude View sounds plausible, I will argue that it is false. In the next section, I will point out the problems of this view.

3. Problems of The Reactive Attitude View

The Reactive Attitude View has two general problems: the wrong kind of reason problem and circularity. In this section, I will explain these problems. Because Fitting Attitude Theories of Wrongness are the most plausible theories for the Reactive Attitude View, I will focus on them. Nevertheless, the wrong kind of reason problem and the circularity problem apply to the other versions of the Reactive Attitude View.

3.1 The Wrong Kind of Reason Problem

The first problem of the Reactive Attitude View is the wrong kind of reason problem. It is commonly accepted (Rabinowicz and Rønnow-Rasmussen, 2004; Jacobson, 2011; Schroeder, 2012) that some factors constitute *reasons for holding* certain reactive attitudes, but they do not constitute *the right kind of reasons for holding* (i.e., the reasons for the *appropriateness* of holding) these reactive attitudes. In other words, a reason for holding an attitude need not bear on the appropriateness of holding the attitude. For example, Customer may have a prudential reason to withhold her anger at Waiter who was rude to her, but that reason does not bear on whether it is appropriate for her to feel

angry at him. This is because the appropriateness of holding an attitude supervenes on the past action alone, not future considerations. Alternatively, suppose a stranger promises to give you ten dollars for feeling gratitude at her punching you. Because of the incentive, you have a prudential reason to feel gratitude towards her, but that reason is not the right kind of reason for you to feel gratitude. Hence, the underlying idea of the wrong kind of reason problem is that certain reasons are of the wrong kind to hold an attitude, which challenges the advocates of the Reactive Attitude View to distinguish the right kind of reasons from the wrong kind. Put otherwise, these theorists need to clarify the nature of the reasons that are relevant for the appropriateness of holding reactive attitudes.

Some might attempt to solve the wrong kind of reason problem by distinguishing historic reasons from instrumental reasons. Clearly, this strategy can rule out many prudential reasons, since incentives and well-being fall into the category of instrumental reasons. Yet, some reasons are not clearly instrumental reasons, though they are the wrong kind of reasons. For instance, etiquette is a reason of the wrong kind for holding an attitude, but it is not clearly an instrumental reason for the reacting people or target agent. More importantly, there is substantive disagreement over whether a reason is of the right kind for holding an attitude. For instance, a moral obligation to a target agent (e.g., obligation to one's family, obligation to compensate) is a reason for holding an attitude towards her, but it is unclear whether it is a reason of the right kind (one that bears on the appropriateness of holding the attitude). Thus, since the Reactive Attitude View builds upon the notion of fitting reactive attitude, proponents of this account have the burden of clarifying the reasons for fitting reactive attitudes (or the right kind of reasons for holding an attitude).

Numerous proposals have been put forward by fitting attitude theorists.⁵² Here, I will examine three proposals only, advocated by Gibbard, Skorupski, and Darwall respectively. I examine their proposals only because they are the fitting attitude theorists of wrongness I discuss. Gibbard appeals to Derek Parfit's distinction between object-given and state-given reasons.⁵³ Object-given reasons are provided by the facts about the *object* (in our case, the target agent of reactive attitudes) and bear on whether the *object* towards whom it is fitting or proper for a reacting person to hold an attitude. In contrast, state-given reasons are provided by the facts about the *state* of a reacting person having certain reactive attitudes towards the object (i.e., the target agent) and bear on whether the *state* is good. Parfit's idea is that reasons of the right kind track whether the target agent is a fitting *object* of the relevant reactive attitudes and thereby are *object-given* reasons, whereas reasons of the wrong kind track whether the *state* of holding the relevant reactive attitudes is good and thereby are *state-given* reasons. Both object- and state-given reasons can be either intrinsic or instrumental.⁵⁴

Parfit's proposal appears powerful in paradigmatic cases of the wrong kind of reason. For instance, in the cases where a stranger promises to give you ten dollars for feeling gratitude at her punching you, you have a reason to feel gratitude to her. Since this reason appears to be a reason for you to be in the state of feeling gratitude, it seems

⁵² Olson, 2004; Hieronymi, 2005; Schroeder, 2007; Way, 2012.

⁵³ Personal communication.

⁵⁴ "Of our reasons to have some desire, some are provided by facts about this desire's object. These reasons we can call *object-given*. We can have such reasons to want something either for its own sake, or for the sake of its effects. . . . Other reasons to want something are provided by facts, not about what we want, but about *having* this desire. These reasons we can call *state-given*. Such reasons can also be either intrinsic or instrumental." Parfit, 2001, pp. 21-22, emphasis original.

that Parfit can reasonably conclude that it is a state-given reason and thereby a wrong kind of reason.

However, the state-/object-given distinction is not clear in the first place. For example, not only does the stranger's promise appear to be a state-given reason, it also appears to be an object-given reason because it is about the stranger. Recall that both intrinsic reasons and instrumental reasons for having an attitude towards an object count as object-given reasons for Parfit, and the stranger's promise is precisely an instrumental reason to feel gratitude to her. Thus, the object-/state-given reasons distinction is blurred.

Not only does the problem of the distinction make it difficult to identify the wrong kind of reason, it also makes it difficult to identify the right kind of reason. For example, it is appropriate for Girl to feel regretful over leaving her loving boyfriend, Guy. Their good relation constitutes a reason for Girl to feel regretful, and it is of the right kind of reason. However, their good relation is both an object-given reason in the sense that it is about Guy (an object of an attitude), and a state-given reason in the sense that it is about whether it is fitting for Girl to be in the state of feeling remorseful. Hence, Parfit needs a better way to distinguish the right kind of reason from the wrong kind.⁵⁵

Furthermore, even if Parfit's distinction between object-/state-given reasons is clear, some wrong kind of reasons are not obviously state-given. For instance, if Agent fails to act rightly because she is coerced, then she has a reason not to feel guilty. As Jacobson (2011) points out, the fact that Agent is coerced does not explicitly refer to any state of having an attitude, but refers to an excuse from being responsible for an action.

⁵⁵ Rabinowicz and Rønnow-Rasmussen (2004), Coleman and Sarch (2012) also critique Parfit's distinction to be unsatisfactory.

Yet arguably it is a wrong kind of reason for her not to feel guilty because the fact that she is coerced does not bear upon the object of guilt (whether such actions are of the sort over which one should feel guilty). Rather, it bears upon the state of her feeling guilty (whether the state of her feeling guilty is adequate).

So far, we have seen that Parfit's attempt fails to solve the wrong kind of reason problem. Next, let us look at Skorupski's proposal. Skorupski attempts to solve the wrong kind of reason problem by distinguishing "reasons to believe or feel on the one hand and reasons to bring it about that one believes or feels on the other."⁵⁶ He claims that this distinction is well-grounded because the former reasons are epistemic reasons (reasons to believe) or evaluative reasons (reasons to have feelings), whereas the latter reasons are practical reasons (reasons to act). If it is legitimate to make this distinction, then he can tell you whether a reason is of the right kind by checking whether it is an evaluative reason. On this proposal, the stranger's promise to give you ten dollars is a practical reason (a reason to bring it about that you feel gratitude to her), not an evaluative reason (a reason for you to feel so). Hence, the promise is a wrong kind of reason.

This proposal is quite attractive since, compared to Parfit's proposal, it better captures the nature of the wrong kind of reason. That is, a reason is a wrong kind of reason for having an attitude because it is not really a reason to have the attitude, but a reason to bring about the case that one has the attitude. The problem of this proposal is that if we have a reason to bring it about that we have an attitude, it follows that this reason gives us a further reason to have the attitude. In particular, if the stranger's promise gives you a reason to bring it about that you feel gratitude, then it is coherent to

⁵⁶ Skorupski, 2007.

infer that you also have a reason to feel gratitude. The reason to feel gratitude, which is derived from a reason to bring about the attitude, may not be a good one, but it nevertheless is a reason to have the attitude.

Anticipating my objection, Skorupski responds that it is *ad hoc* to distinguish good reasons from bad reasons to have an attitude.⁵⁷ He argues that distinguishing between reasons to have an attitude from reasons to bring about an attitude equally explains the cases as distinguishing good reasons from bad reasons for an attitude does. He further argues that the former distinction is more natural than the latter.

I agree with Skorupski that the two distinctions are equivalent in terms of explanatory power, for they both explain why some reasons are of the right kind while others are not. However, I contend that the distinction between good and bad reasons to have an attitude is preferable to the distinction between reasons to have an attitude and reasons to bring about an attitude.

First of all, since a reason for having an attitude is roughly a consideration counting in favor of having the attitude, it is plausible that some reasons are bad reasons for having an attitude. It will be surprising if all reasons for having an attitude are good ones. Second, the reason for bringing about an attitude is a particular sort of practical reason: it presupposes a standard to determine what makes it worthwhile bringing about an attitude. Without providing the standard, Skorupski's proposal fares no better than Parfit's. Skorupski is aware of the presupposition, but complains that it does not beg the question and claims that the burden of proof is on the opponents of the Reactive Attitude View. This reply is unsatisfactory, for a good solution to the wrong kind of reason

⁵⁷ Skorupski, 2007.

problem has to be complete, providing both a rationale for drawing the distinction between the right and wrong kind of reasons and a procedure for drawing it. If all Skorupski succeeds in doing is to draw a dubious distinction without offering rationale for it, then his proposal does not solve the problem.⁵⁸

The failure of Parfit's and Skorupski's proposals does not amount to the total failure of the Reactive Attitude View. Darwall (2006) proposes a psychological test to track the right kind of reason. In what follows, I will examine his proposal.

Darwall holds that a consideration is a reason of the wrong kind for holding a reactive attitude if and only if it is psychologically impossible for a reacting agent to hold the attitude after deliberating from the consideration. In particular, reason of the right kind

must be a fact about or feature of some object, appropriate consideration of which could provide *someone's reason* for a warranted attitude of that kind towards it. It must be something on the basis of which someone could (and appropriately would) come to hold the attitude as a conclusion of a process of considering (deliberating about) whether to do so.⁵⁹

⁵⁸ It makes sense to require the opponents of the Reactive Attitude View to do the same thing. One natural answer is that we can draw the distinction between good and bad reasons (or right and wrong reasons) by checking whether the reasons at stake are provided by the moral properties of the object. For example, we shall examine whether the object is morally good or morally right.

⁵⁹ Darwall, 2006, p. 16, emphasis original.

Hence, for Darwall, we could ask this question to know whether a reason is of the right kind for having an attitude: “Is it psychologically possible for anyone to come to hold the attitude after considering this reason?”

As Coleman and Sarch (2012) points out, Darwall’s test makes the appropriateness of an attitude depend upon human psychology. This implies that the appropriateness is not solely a matter of what the target agent has done, but the whole spectrum of reacting people’s possible responses. This appears implausible, for a fitting reactive attitude should only fit the target agent’s particular action in the relevant circumstances, so reacting people’s possible responses are irrelevant.

Darwall might respond that although reacting people’s possible responses are not conceptually relevant to fitting attitudes, they are psychologically relevant. Specifically, it just happens that we are all psychologically constructed such that we cannot come to have an attitude if we merely consider a wrong kind of reason to have the attitude.

This response works in some cases, since, for instance, it is not easy for you to feel gratitude to a stranger just because she promises to give you ten dollars if you feel so. However, Darwall underestimates our capacity to bring about an attitude. If we really want to have an attitude, we can work on having it and might succeed in having it. In the previous case, you might try to imagine that the stranger has done something nice to you and make yourself believe that this really happens. It is psychologically possible for you to succeed in coming to have the belief, and it is also psychologically possible for you to feel gratitude to the stranger as a result. Hence, Darwall’s proposal does not work either.

We have seen that Parfit’s proposal, Skorupski’s proposal, and Darwall’s proposal each fails to solve the wrong kind of reason problem. Although the proposals examined

above all fail, it does not follow that the Reactive Attitude View fails. In any case, other proposals exist for exponents of this view to solve the wrong kind of reason problem, and one of them might eventually turn out successful. However, given that the proposals discussed above are leading ones, their failure seems to suggest that Reactive Attitude View is unable to solve the problem.

The wrong kind of reason problem may turn out to be a technical problem in the end. Even if this problem is solved, the Reactive Attitude View still faces a more serious problem: the circularity problem.

3.2 Circularity

The Reactive Attitude View, I claim, is circular. For example, when moral wrongness is defined in terms of fitting reactive attitudes, such attitudes need to be independent of moral wrongness. However, the Reactive Attitude View often sneaks in the notion of moral wrongness when it defines fitting reactive attitudes. As a result, this account is circular.

In my view, all fitting attitude theorists of wrongness have the circularity problem, including the contemporary versions. In general, the notion of blameworthiness presupposes the notion of wrongness. For example, it is natural to say that “You are blameworthy because you did something wrong,” or “You are blameworthy because you wronged her.” These expressions suggest that wrongness underlies blameworthiness. Accordingly, without the appeal to wrongness, blameworthiness seems difficult to spell out. This idea makes it plausible that Fitting Attitude Theories of Wrongness are all circular.

The challenge of circularity to Fitting Attitude Theories of Wrongness is not new. In fact, Broad (1928) and Ross (1939) both hold that rightness is a kind of fittingness to the situation, but they deny that rightness can be defined in terms of fittingness. Ross insightfully says that:

[W]e can begin to define moral rightness, because we can say it is a form of suitability; but we cannot complete the definition, since if we ask what kind of suitability it is we can only say that it is the kind of suitability that is rightness.⁶⁰

Despite the familiarity of the circularity challenge, it is important to see why the contemporary fitting attitude theorists are all subject to the problem. The reason is that circularity can be avoided as long these theorists pick out a single notion that does not presuppose moral wrongness.

Let us now examine Darwall's, Skorupski's, and Gibbard's accounts respectively. Darwall proposes that fitting reactive attitudes are defined in terms of second-personal reasons, "a distinctive kind of (normative) reason for acting,"⁶¹ a reason made on one's will and purportedly given by an authority's demand. For people to enter second-personal relations, the kind of relation where one's demand generates a second-personal reason for the other to act, they have to be free and rational agents.⁶² In Darwall's favorite example,

⁶⁰ Ross, 1930, p. 55.

⁶¹ Darwall, 2006, p. 4.

⁶² Again, Darwall thinks that one can enter second-personal relation with oneself. Here I talk about two parties, for the sake of simplicity.

you step on my foot without any excuse, then you have a second-personal reason to remove your foot from the top of mine because (i) I have the authority to issue a demand to you, (ii) you are competent to understand and follow my demand, and (iii) I demand that you remove your foot.⁶³ When you step on my foot without any excuse, it is appropriate for me to feel resentful towards you given what you've done, and the appropriateness comes from your second-personal reason to remove your foot. In short, Darwall takes second-personal reasons, which are grounded by second-personal relations, to explain fitting reactive attitudes.

It may seem that Darwall's notion of fitting reactive attitudes does not appeal to moral wrongness, for the core about second-personal reasons is merely the competence of the agents involved. In the case where you step on my foot, both you and I are free and rational beings in the moral community, so we both have the authority to address each other and the competence to acknowledge the validity of the demands each other makes. Nothing about our competence refers back to moral wrongness. Hence, Darwall's account does not appear circular at a first glance.

I claim, however, that Darwall's account is circular because, for a second-personal reason to hold, the corresponding demand must be valid, which entails that a wrong has been done. For example, in the foot-stepping case, you have a second-personal reason to remove your foot crucially because my demand that you do so is valid. Yet the reason that my demand is valid is just that you have wronged me in stepping on my foot without any excuse. Hence, Darwall is right that it is appropriate for me to feel resentful towards you given what you have done, but the ultimate reason that my resentment is

⁶³ Darwall, 2006, pp. 5-8.

appropriate is not that you have a second-personal reason to remove your foot, nor that we are both competent to enter second-personal relations. Rather, the reason simply is that you have acted wrongly. Therefore, when Darwall explains fitting reactive attitudes in terms of second-personal reasons, he implicitly brings in moral wrongness.

Darwall might reply that he does not need wrongness to explain the validity of a demand. To say that a demand is valid, he might say, is just to say that the target agent had no excuse in performing the relevant action.⁶⁴ Thus, in the case you step on my foot, my demand that you do not step on my foot is valid simply because you had no excuse in doing so. In other words, you had no explanations to exculpate yourself from stepping on my foot, which makes my demand valid.

To examine this reply, we have to understand what excuses involve in the first place. When there is an excuse, there are conditions where an agent's agency is undermined: she is not informed of the relevant facts, she is not capable of making autonomous decisions (e.g., too young, mentally sick), or her action is not voluntary (e.g., under duress). We examine whether these conditions exist precisely because we want to find out whether the agent needs to be exculpated from *wrongdoing*. For all we know, cases of excuses are just cases having extenuating conditions. In other words, the notion of excuse presupposes wrongdoing, for the question of accountability for wrongdoing just doesn't arise in the context of right actions.

Of course, wrongdoing is relative to a set of standards about right and wrong, which includes both moral and nonmoral standards. Thus, wrongdoing does not necessarily refer to *moral* wrongdoing. Along this line, Darwall might argue that all he

⁶⁴ Thanks to Robert Johnson for this suggestion.

appeals to is a generic notion of wrongness, not moral wrongness in particular, and hence his account is not circular.

The problem with this line is that it implies that every wrong is a moral wrong. Specifically, Fitting Attitude Theories of Wrongness explain moral wrongness in terms of blameworthiness. If Darwall explains blameworthiness in terms of generic wrongness, then an agent would be regarded to have acted morally wrong when her action is non-morally wrong. This is because a non-moral wrong could make an agent blameworthy (e.g., the failure to help one's soccer team win a friendly), and on Fitting Attitude Theories of Wrongness the agent would have acted morally wrongly when she had not excuse. Since it is implausible to regard every wrong as a moral wrong, this line fails. Thus, the appeal to excuse does not avoid the problem of circularity.

Alternatively, Darwall might reply that he can appeal to reasons to explain the validity of a demand.⁶⁵ The idea is that a demand is valid if and only if there are reasons to justify it, where the reasons may be rights, obligations, well-being, virtues, consequences, etc. Take the foot-stepping case again. According to this line, my bodily right to my foot justifies my demand that you remove your foot, which implies that my demand is valid and my resentment to you is fitting. Moreover, the appeal to moral wrong is unnecessary.

I agree that various moral reasons justify the validity of a demand. However, the circularity problem reappears because all justifying reasons for a valid demand constitute reasons that the target agent acts *morally wrongly*. For example, my bodily right to my foot not only justifies my demand that you remove your foot, but it also constitutes the

⁶⁵ Thanks to Robert Johnson for this suggestion.

reason that you act wrongly when you step on my foot without any excuse. It's just a further specification of the reason you act wrongly. In other words, it is not distinct from the reason of moral wrongdoing, and thus the appeal to the notion of moral wrongness is not avoided. Accordingly, Darwall can't appeal to justifying reasons to solve the circularity problem.

Darwall could reply in two ways. First, he might be skeptical whether moral wrongdoing provides reasons at all. If so, he can deny that moral wrongdoing provides a reason for a second-personal demand to be valid. This reply is not successful. One reason is that Darwall (2010) has explicitly argued that moral wrongdoing has reason-conferring powers. Another reason is that folk psychology suggests that moral wrongdoing provides reasons. For example, it is natural to say that "It's wrong to break a promise/kill a person/... So, you shouldn't do so." Moreover, it is implausible that the folks are massively mistaken.

Second, Darwall might contend that although moral wrongdoing provides reasons, it is completely reduced to justifying reasons. This idea is that moral wrongness is nothing above and beyond various justifying reasons, which reasons directly explain why a demand is valid. If this is correct, then the appeal to moral wrongness is unnecessary.

This reply deserves scrutiny. First of all, moral wrongness is not merely a set of reasons for action. For example, for a particular action, some reasons favor it, some are against it, while others are not relevant. Whether the action is morally wrong bears upon which reasons are relevant and how weighty these reasons are, while there are no universal principles to answer these questions. Hence, it is not plausible that moral wrongness is nothing but some reasons for actions.

Second, there are many reasons for actions, so we have to distinguish the ones concerning the validity of a demand from the rest of the reasons. This brings us back to the wrong kind of reason problem. In light of the discussions in Section 3.1, I conclude that this problem cannot be solved, and hence it does not help Darwall to appeal to justifying reasons.

It might seem promising for Darwall to bite the bullet, allow that his account to be circular, but maintain that his account is not viciously circular. In particular, he might argue that the talk of second-personal reasons illuminates when reactive attitudes are fitting without the direct appeal to the mysterious notion of wrongness. Perhaps in the case you step on my foot, it is more intelligible to say that *my resentment towards you is fitting because you have a second-personal reason not to step on my foot* instead of saying that *my resentment towards you is fitting because you have wronged me*. The underlying idea might be that the talk of reasons is easier to understand than the talk of wrongness.

I cannot see why the appeal to second-personal reasons is more intelligible than the appeal to wrongness. The notions of reason and wrongness are both abstract, and the notion of second-personal reason involves complex relation between the target agent and the reacting person. In contrast, the notion of wrongness might be easier to grasp.

Anyway, even if the notion of a second-personal reason is easier to grasp than that of moral wrongness, Darwall would abandon Reactive Attitude View if he bit the bullet. This is because one has to maintain the conceptual priority of fitting reactive attitudes to hold that account, and explaining blameworthiness in terms of wrongness amounts to admitting that blameworthiness is not conceptually prior to wrongness.

Let's now consider Skorupski's account instead. He intends to define blameworthiness via blame without the notion of moral wrongness. Specifically, he defines blame by the actions it characteristically disposes, which he takes to be "withdrawal of recognition."⁶⁶ "To withdraw, withhold, or refuse recognition," he explains, "is to cut off relations or refuse to enter into them, to exclude, in however partial and temporary a way; in more extreme cases it leads to ostracism, casting out, outlawing."⁶⁷ Moreover, the recognition blame disposes is moral recognition: "recognition of the other as a *responsible moral* agent, i.e. one who possesses the subjectivity and moral insight of a moral agent, and can be relied on, and if necessary called on, to act on it."⁶⁸ Based on this specific notion of blame, Skorupski defines blameworthiness as the sentiment that is both "the appropriate response to certain representative examples of belief and action" (i.e., worthiness) and "prompts a characteristic-type of action, namely, withdrawal of recognition" (i.e., blame).⁶⁹

I claim that Skorupski's account is incomplete. First of all, it is inaccurate to construe blame merely as a behavior disposition. Blame is essentially dependent upon explanatory the reason of blame. Without incorporating the reason, we can hardly distinguish blame from contempt, for example. This is because both blame and contempt dispose us to "cut off relations" with the target agent. According to the Oxford English

⁶⁶ Skorupski, 2010b, p. 294.

⁶⁷ Skorupski, 2010b, p. 294.

⁶⁸ Skorupski, 2010b, p. 374, emphasis original.

⁶⁹ Skorupski, 2010, p. 295.

Dictionary, the noun blame means “responsibility for a fault or wrong.”⁷⁰ This definition makes the explanatory reason of blame explicit, which suggests that the reason is central to the meaning of blame. This is because, according to the definition, blame is not responsibility at the abstract level, but responsibility for a *fault* or *wrong*.

It will be hasty to reject Skorupski’s account because of the incompleteness of his account, since he does discuss how to distinguish blame from close emotions such as horror, disgust, contempt, and disdain merely by behavioral disposition. The key, he says, “is that the rejection to which these [i.e., horror, disgust, contempt, and disdain] dispose does not, as does that involved in blame, inherently posit possibility of restoration of standing.”⁷¹ In other words, blame is distinct from emotions such as horror and disgust in that it essentially admits the possibility for the target agent to regain (moral) recognition from other agents. The distinction seems to make sense, since we normally would stop blaming an agent whom we believe to have rectified her wrong, while it seems less clear whether we would normally stop feeling horrified of, or contempt for, an agent when we saw her act differently from what she used to.

I contend that Skorupski’s strategy somewhat helps distinguish blame from horror, but does not succeed across the board. According to him, we arguably have horror feelings because we “treat the transgressive item [e.g., the target agent’s sin] as a polluting impurity or uncleanness.”⁷² Even if we stop feeling horrified at the object, he

⁷⁰ http://oxforddictionaries.com/definition/american_english/blame?region=us&q=blame

⁷¹ Skorupski, 2010b, pp. 376-377.

⁷² Skorupski, 2010b, p. 377.

continues, the feelings remain “latent and can still influence ideas of the moral.”⁷³ What Skorupski has in mind is Christianity’s doctrine of sin. It is plausible that when a devoted Christian somehow stops believing in Christianity, she will still have some residual ideas about sin, which affect her understanding of morality. However, that case does not show that horror inherently excludes the possibility of restoration. This is because the impact of horror is usually not as sweeping as religious horror: in many cases (e.g., horror at ghosts, revenge), once we think that the object does not exist, or exists but is no longer horrible, we will not feel horrified any more. Thus, the talk of the possibility of restoration does not help distinguish blame from horror in many cases.

The distinction is more blurred when it comes to blame versus contempt. If we hold an agent in contempt, then we think that she deserves no respect at all and perhaps even deserves disrespect. Our thinking so in many cases is due to our impressions of or the beliefs about the things she has said and the way she behaves, while we do not respect people saying those things and behaving in her way. But if she behaves differently (e.g., wisely, intelligently) and we notice that, then after a while, as open-minded people, we may stop holding her in contempt and start to respect her. I don’t see why contempt excludes the possibility of restoration of standing and thereby I do not find Skorupski’s strategy to distinguish contempt from blame convincing.

In my view, the key difference between blame and contempt, blame and horror, etc. is that we hold these feelings for different *reasons*. In particular, we have the blame-sentiment because we think the target agent has acted wrongly without any excuses, while we don’t have horror feelings or hold anyone in contempt for this reason. Without

⁷³ Skorupski, 2010b, p. 377.

the talk of the reasons, merely the talk of behavioral disposition does not help distinguish blame from many other emotions.

The second problem of Skorupski's account is that even if his notion of blame is tenable, his notion of blameworthiness is not. This is because he still explains blameworthiness as certain "appropriate response," which needs unpacking. The appropriateness of blame must depend upon reasons, since Skorupski believes that blameworthiness "is concerned with the agent's belief-relative reasons," namely, the reasons that are relative to the agent's factual beliefs.⁷⁴ In particular, he accepts the following claim:

(2) If x 's a -ing is blameworthy, then if x 's factual beliefs were true there would be reason for x not to a .⁷⁵

Skorupski gives this example: if you cure a patient by administering a drug, but you mistakenly believe the drug to be a poison and intend to poison the patient, then you are blameworthy for what you have done.⁷⁶ You are blameworthy, according to Skorupski, "just because you believed yourself to be poisoning the patient."⁷⁷

⁷⁴ Skorupski, 2010b, p. 295.

⁷⁵ Skorupski, 2010b, p. 296.

⁷⁶ Skorupski, 2010b, p. 296. The example comes from Parfit (forthcoming), Section 21.

⁷⁷ Skorupski, 2010b, p. 296.

I argued in Chapter 3 that an agent is blameworthy if and only if she inexcusably acts wrongly. Hence, I disagree with Skorupski's judgment. Since what I need to show is that Skorupski's notion of blameworthiness presupposes moral wrongness, we can bypass the disagreement.

Skorupski's notion of blameworthiness presupposes the notion of moral wrongness. To see this, we need to notice that Skorupski has not adequately explained your purported blameworthiness in this case. The content of your belief does not adequately explain why you seem blameworthy; rather, *it being wrong to administer the drug if your belief were true* is the right explanation. Put another way, you are blameworthy for administering the drug to the patient not just because there would be reason for you not to do so were your belief true, but because, more precisely, it would be *wrong* for you to do so were your belief true. Without the appeal to wrongness, it is difficult to figure out why your holding a false belief while administering the drug makes you seem blameworthy.

In virtue of the explanation Skorupski provides for the case above, I take him to hold

(2*) x 's a -ing is blameworthy in virtue of the fact that were x 's factual beliefs true, there would be reason for x not to a .

We can express the meaning of (2*) by simply saying:

(2#) x 's a -ing is blameworthy in virtue of the fact that were x 's factual beliefs true, it would be wrong for x not to a .

(2#) is more precise than (2*), for there would be reason for x not to a precisely because it would be wrong for x not to a . If I am right, then wrongness grounds or explains blameworthiness.

Skorupski might respond in two ways. First, he might argue that (2*) and (2#) are not semantically equivalent. A straightforward method is to argue that an action being wrong does not provide any reason against performing the action. This line is not open to Darwall, but it is open to Skorupski.

However, this line does not fit the way the folks talk about reasons for actions. If the folks are wrong in thinking that wrongdoing provides reasons against actions, they have to be massively mistaken. If so, then arguably the talk of wrongdoing seems unnecessary and can be eliminated. But I don't think that moral talk is meaningful in this way. Hence, I maintain that if an action is morally wrong, then its moral wrongness provides a reason not to perform the action.

Another way for Skorupski to argue that (2*) and (2#) are not semantically equivalent is to claim that an action's being wrong is not the only reason for an agent not to perform the action, were the agent's factual beliefs true. Perhaps an agent has a reason not to perform an action despite the action being right because she would have exhibited a bad character trait in performing it. Suppose Batman fails to save the people in Gotham City because he is busy developing his business and he believes his business to be more important than these people. Further suppose that Batman does not act wrongly in doing so (because it is his supererogation, not moral obligation). It looks as if Batman has a reason to save the people in Gotham City, but if his belief were true, he no longer has this reason. Hence, this case fails to help Skorupski. We can generalize this case. In any case

where an agent acts rightly but is criticizable for something (e.g., bad traits, bad consequences), she either has a belief that justifies her action or she is unaware of the criticizable thing. If she has such belief, then were her belief true, her action would be justified and thus she would have no reason to refrain from the action. If she is unaware of the criticizable thing, then she believes herself to act rightly. And if this belief were true, she would not have a reason to refrain from the action either. Thus, I can't see how (2*) and (2#) can be semantically not equivalent.

The second strategy Skorupski can use is accepting (2*) and (2#) to be semantically equivalent but arguing that the notion of reason is more primitive than the notion of moral wrongness. In fact, it is likely that he will adopt this strategy, since one of his central claims in *The Domain of Reasons* (2010b) is that “the fundamental normative concept is that of a reason relation” and reason is a primitive notion.⁷⁸ Even if Skorupski is right about the primitiveness of reason, accepting the semantic equivalence between (2#) and (2*) entails that both reason and moral wrongness underlie blameworthiness. Certainly, Skorupski might claim that reason underlies blameworthiness in a more robust sense. Nevertheless, what we need is just that moral wrongness underlies blameworthiness. Hence, the strategy of arguing for reason being primitive does not work. Since neither strategies work for Skorupski, I conclude that his account of blameworthiness is circular.

Let us now take a look at Gibbard's account. Recall that he attempts to reduce subjective wrongness (roughly, wrongness relevant to the agent's epistemically justified beliefs) to the warrant of guilt and resentment. After conversion, his account becomes: an

⁷⁸ Skorupski, 2012a, p. 174; Skorupski 2012b, p. 210.

action is objectively morally wrong if and only if, given that both the agent and the reacting people are fully informed, both the agent's feeling guilty for having done it is warranted and the reacting people's feeling resentful towards her for having done it is warranted.

For Gibbard's account of objective wrongness to work, he has to define the warrant of guilt and resentment independently of objective wrongness. However, it seems that no plausible standard of the warrant of reactive attitudes is possible without presupposing objective wrongness. Specifically, I claim that a fully informed agent's feeling guilty over performing an action is warranted if and only if, and *in virtue of the fact that*, her action is (objectively) morally wrong and she had no excuse. For example, feeling guilty for Grandson over not helping Grandma with her groceries is warranted if and only if his failure to help her is morally wrong and he had no excuse. In contrast, whether his failure to help Grandma is morally wrong does *not* depend upon whether his guilt about the failure is warranted, given that he is fully informed. Similarly, feeling resentful for fully informed people towards an agent over her performing an action is warranted if and only if, and *in virtue of the fact that*, the action is (objectively) morally wrong and the agent had no excuse. Simply put, there is no stand-alone sense of whether the feelings of guilt or resentment is warranted apart from the notion of (objective) moral wrongness.

Gibbard is certainly aware of this concern. He regards warrant as a thin sense of ought or a primitive ought. This ought is neither moral nor nonmoral; rather, it is all things considered, "what it makes most sense to do [or to believe or feel], everything

considered.”⁷⁹ Alternatively, he explains this ought as conclusive reasons.⁸⁰ In his view, since the warrant of guilt and resentment does not presuppose objective wrongness, his account of objective wrongness is not circular. His explanation of the primitive ought in *Reconciling Our Aims* is very clear:

I am contrasting, then, oughts in general and moral oughts. Narrowly moral questions of right and wrong I’m treating as at base questions of what moral sentiments we ought to have and act from. Questions in the broader class of oughts in general we call *normative* questions. These include questions of what a person ought to do all things considered. They include epistemological questions of what we ought to believe. And they include questions of how we ought to feel about things. These, I am saying, are all, in a broad, extended sense, planning questions; they are questions of what to do, to think, and to feel. Moral questions are planning questions of a particular kind, questions of how to feel about things, where the feelings in question are the moral sentiments.⁸¹

According to Gibbard, the question for Grandson is what he ought to feel, which amounts to a plan over what feelings to have. The correct answer is guilt. And if Grandson does not plan to feel guilty, Gibbard will propose that Grandson has to improve

⁷⁹ Gibbard, 2008b, p. 176.

⁸⁰ Gibbard, 2013, p. 15, footnote 26.

⁸¹ Gibbard, 2008a, p. 17, emphasis original.

his sensibilities.⁸² But two questions follow from this proposal: (1) Is it intelligible to understand the primitive ought as planning? (2) Why does Grandson need to improve his sensibilities if he doesn't plan to feel guilty? It seems that the real reason is that not helping Grandma is objectively morally wrong (and he lacks full excuses).

Gibbard finds no difficulty in positively answering the first question. When challenged by John Broome how to understand hypothetical planning, Gibbard suggests that when we plan to feel or do something, we “just reject ruling it out by preference” in a “hypothetical frame of mind.”⁸³ Thus, when Grandson judges that he ought to feel guilty, he makes the hypothetical plan to feel guilty, which is explained in terms of rejecting ruling out preferring any alternative to the plan to feel guilty.

This explanation is not very satisfying. In the first place, as Broome points out, the question “what ought I to feel” and the question “what shall I feel” are two distinct questions. The answer to the one need not be the answer to the other. Furthermore, even if the answer to both questions converge, we make a plan not merely because we prefer it to other plans, but also because the plan is normative for us to carry out. Thus, if Grandson plans to feel guilty, it is not merely because he prefers guilt to other feelings, but more importantly because he *ought* to feel guilty. This problem is connected to the question why Grandson has to improve his sensibilities if he doesn't feel guilty, so I will turn to that question now.

A natural thought is that objective moral wrongness figures in the explanation of what sensibilities to have. If we take this line, then the primitive ought explains what plan

⁸² Gibbard, forthcoming.

⁸³ Broome, 2008. Gibbard, 200b, p. 177.

to have, instead of the other way around. Gibbard obviously rejects this line. He worries that anyone taking this line either can't explain what 'ought' means, or "their answer involves strange and incredible things like non-natural properties."⁸⁴ As an alternative, he proposes that the feature of an action explains what action, belief, or feelings to have. Specifically, when Broome challenges that we have a plan only if we believe that we ought to carry out the plan, Gibbard replies that when we accept/reject a plan, we just need a belief that an action has certain feature, not any belief about ought:

I reject any plan to hit my thumb with a hammer, and to do this, why would I need to believe, even implicitly, that I 'ought' not to do so or that it would be a 'bad' thing to do? I just need the belief that it would hurt like hell, along with the absence of any countervailing beliefs...⁸⁵

Gibbard is right that sometimes we need not think about what we ought to do when we accept a plan. However, the features of a plan bear upon whether we ought to accept it. For example, the idea that Gibbard ought not to hit his thumb with a hammer is justified by the idea that doing so would hurt him. He need not explicitly think about whether he ought to perform the action, but when he thinks about the consequences of the action, he is indirectly thinking about the deontic status of the action. Moreover, sometimes we accept or reject a plan precisely because we believe that we ought or ought not to carry out the plan. In other words, sometimes we perform an action just because we feel its

⁸⁴ Gibbard, 2008a, p. 20.

⁸⁵ Gibbard, 2008b, p. 178.

normative force. “I have to do it.” “It’s my obligation.” These thoughts do play a role in our daily plans. Therefore, what we ought to do/believe/feel explains what plans to have regarding actions/beliefs/feelings.

So far, I have argued that the primitive ought explains the warrant of guilt and resentment, but I have not established that objective moral wrongness explains that warrant. To make the case that Gibbard’s account is circular, I have to fill the gap between objective moral wrongness and the primitive ought. Recall that the primitive ought is an all-things-considered ought, so it bears upon every kind of reason—epistemic, moral, prudential, aesthetic, etc. I argue that when we face the question whether we all-things-considered ought to feel guilty or resentful towards an agent, our question becomes whether we morally ought to feel guilty or resentful towards the agent. The reason is that when it comes to the feelings of guilt and resentment, moral considerations outweigh all the nonmoral ones. So when we take every consideration into account, whether we ought to feel guilty or resentful turns into the question of whether we *morally* ought to feel so. Thus, Gibbard’s account of objective moral wrongness ultimately presupposes objective moral wrongness, which makes it circular.

Thus far, I have argued that Darwall’s, Skorupski’s, and Gibbard’s accounts of wrongness are all subject to the circularity problem. I conclude that the Reactive Attitude View is false because its most plausible version—Fitting Attitude Theories of Wrongness—is subject to the wrong kind of reason problem and the circularity problem.

4. Conclusion

In this chapter, I have argued against the Reactive Attitude View. On the Reactive Attitude View, moral wrongness is defined, at least partly, by the normative status of our having (or expressing) core negative reactive attitudes, but not vice-versa. I argued that this view is false because it defines moral wrongness in terms of the wrong kind of reason and it is circular.

In the next chapter, I will reject the Independence View and the Interdependence View. The former view maintains that moral wrongness is not defined in terms of the normative status of our having/expressing core negative reactive attitudes, nor is such normative status defined in terms of moral wrongness. The latter view maintains that moral wrongness and such normative status are defined in terms of each other.

Ch. 5 The Interdependence View and The Independence View

In this chapter, I will argue against the Interdependence View and the Independence View. The Interdependence View holds that moral wrongness is defined by the normative status of our having (or expressing) core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt), and such normative status is also defined in terms of moral wrongness. The independence View holds that neither moral wrongness is defined in terms of such normative status, nor is such normative status defined in terms of moral wrongness. I will first reject the Interdependence View and then reject the Independence View.

1. Against The Interdependence View

1.1 The Interdependence View

The Interdependence View takes the Non-Naturalist Reactive Attitude Approach, so it accepts the Non-Naturalistic Reactive Attitude Thesis. That is, it maintains that moral wrongness is conceptually connected with the normative status of our having (or expressing) core negative reactive attitudes (i.e., blame, resentment, indignation, and guilt). What distinguishes the Interdependence View is the commitment that moral wrongness and such normative status are defined, at least partly, in terms of each other. In brief, this view maintains that the concepts of moral wrongness and such normative status are interdependent. Formally, the Interdependence View claims the following:

Interdependence View:

- (1) It is morally wrong for an agent, S, to perform an action, A, if and only if, were S to have no excuse, then it would be appropriate (permissible, or obligatory) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A.
- (2) The moral wrongness of A *is* in virtue of the appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A.
- (3) The appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A *is* in virtue of the moral wrongness of A.

Claim (1) is the Non-Naturalistic Reactive Attitude Thesis. The Interdependence View is distinctive from the Reactive Attitude View in proposing claims (2) and (3) together.

Because it is more plausible to discuss appropriateness than to discuss permissibility or obligation, hereafter I will discuss *appropriateness* alone. In addition, it is more plausible to discuss the appropriateness of *holding* reactive attitudes than to discuss the appropriateness of *expressing* reactive attitudes, so I will discuss the *appropriateness of holding* reactive attitudes only.

David Wiggins (1991) and, arguably, John McDowell (1985) hold the Interdependence View. Both McDowell and Wiggins claim that the value of an object is in virtue of the object being an appropriate object of our approbation, while the appropriateness of our approbation is in virtue of the object's being good. In Wiggins'

words, “ x is good if and only if x is the sort of thing that calls forth or makes appropriate a certain sentiment of approbation *given the range of propensities that we actually have to respond in this or that way.*”⁸⁶ This formulation is ambiguous between the object *causing* the sentiment and that *meriting* the sentiment. Yet we can see that he favors the latter because he says that a good object “is such as to make a certain sentiment of approbation *appropriate.*”⁸⁷ While McDowell merely extends this view to some value concepts such as the dangerous and the funny, Wiggins extends it to all value concepts (e.g., the bad, beautiful, ugly, brave, just, disgusting, and funny) and to the deontic concept of the right. This is because he maintains that “ x is good/right/beautiful if and only if x is such as to make a certain sentiment of approbation *appropriate.*”⁸⁸

1.2 Motivations for the Interdependence View

Some might be attracted by the Interdependence View because they think that our experiences with deontic concepts are analogous to our color experiences. Both McDowell and Wiggins hold that we cannot understand the notion of red without our experiencing an object looking red, whereas an object looks red precisely because it is red.⁸⁹ If they are right, and deontic concepts are analogous to color concepts, then an

⁸⁶ Wiggins, 1991, p. 206, emphasis original.

⁸⁷ Wiggins, 1991, p. 187.

⁸⁸ Ibid.

⁸⁹ “...an object’s being red is understood as something that obtains in virtue of the object’s being such as (in certain circumstances) to look, precisely, red.” McDowell, 1985, p. 133. “ x is red if and only if x is such as to give, under certain conditions specifiable as normal, a certain visual impression’ naturally raises the question ‘which visual impression?’ And that question attracts the answer ‘an impression as of seeing something red’, which reintroduces *red.*” Wiggins, 1991, p. 189.

action is morally wrong mainly because we have the blame-sentiment towards the agent, and we have the sentiment mainly because the action is morally wrong. Thus, moral wrongness and the normative status of our having/expressing core negative reactive attitudes depends upon each other. In other words, neither moral wrongness nor that normative status is conceptually prior to the other.

1.3 Problem of the Interdependence View

I accept claims (1) and (3), but reject claim (2). Claim (1) states that it is morally wrong for an agent, S, to perform an action, A, if and only if, were S to have no excuse, then it would be appropriate (permissible, or obligatory) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A. Since this claim is the Non-Naturalistic Reactive Attitude Thesis, and I defended it in Chapter 3, I deem claim (1) true. Claim (3) states that the appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A is in virtue of the moral wrongness of A. I deem this claim to be true, but I need not defend it now (I will defend it partly in Section 2.3 and fully in Chapter 6). Rejecting claim (2) suffices to reject the Interdependence View.

Before I address claim (2), I would like to point out an internal problem of the Interdependence View: endorsing both claims (2) and (3) amounts to a circular view. The reason is that claim (2) explains moral wrongness in terms of the appropriateness of core negative reactive attitudes and claim (3) explains such appropriateness in terms of moral wrongness. Hence, when the Interdependence View upholds both claims, it is circular.

Both McDowell and Wiggins are aware that their accounts of value are circular, but they argue that their accounts are correct and the circularity is benign. Although they focus on accounts of value, their points apply to accounts of moral wrongness. McDowell claims that we have to take a no-priority view about value because values “are not brutally there—not there independently of our sensibility—” and “there is no comprehending the right sentiments independently of the concepts of the relevant extra features.”⁹⁰ That is, when we can neither understand value independently of our sensibility nor understand our fitting sensibility independently of value, a circular view is the only option. Wiggins further explains that his account is not a definition of value, but an explanation of value. As he puts it, he “hopes to elucidate the concept of value by displaying it in its actual involvement with the sentiments. One would not, according to [my view], have sufficiently elucidated what value is *without* that detour.”⁹¹ Thus, the purpose is not offering definitions, but offering helpful explanations. Presumably offering such explanations allows circularity.

I don’t agree that the Interdependence View is the correct account we are left with. I will defend the Standard View in the next chapter. More importantly, I don’t agree that a circular explanation of moral wrongness is helpful for us to understand the concept. The reason is that such explanation ultimately explains moral wrongness in terms of moral wrongness, which is not informative at all. Even if this internal problem of circularity is not serious, the Interdependence View is false because claim (2) is false.

⁹⁰ McDowell, 1985, p. 146, p. 160.

⁹¹ Wiggins, 1991, p. 189.

Claim (2) states that the moral wrongness of A is in virtue of the appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A. This claim is central to the Reactive Attitude View, for that view accepts that the appropriateness is conceptually prior to moral wrongness. I rejected the Reactive Attitude View in Chapter 4. If I am right that we cannot non-circularly explain moral wrongness in terms of the appropriateness, then claim (2) is false. As a result, I declare that the Interdependence View is false.

I will turn to the Independence View in the next section.

2. Against The Independence View

2.1 The Independence View

Like the Interdependence View, the Independence View adopts the Non-Naturalist Reactive Attitude Approach. What distinguishes the Independence View from the other views taking the Non-Naturalist Reactive Attitude approach is that it holds that neither moral wrongness is defined in terms of such normative status, nor is such normative status defined in terms of moral wrongness. That is, it takes moral wrongness and that normative status to be conceptually independent of each other. Formally, the Independence View claims the following:

Independence View:

- (1) It is morally wrong for an agent, S, to perform an action, A, if and only if, were S to have no excuse, then it would be appropriate (permissible, or obligatory) for

reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A.

(4) The moral wrongness of A is *not* in virtue of the appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A.

(5) The appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A is *not* in virtue of the moral wrongness of A.

Again, claim (1) is the Non-Naturalistic Reactive Attitude Thesis. The Independence View is distinct in proposing claims (4) and (5) together. Again, because of plausibility, I will discuss the *appropriateness of holding* reactive attitudes only.

I do not know of anyone endorsing the Independence View, but it is helpful to explore this view for the completeness of all possible views to the problem of the conceptual relationship between moral wrongness and reactive attitudes.

2.2 Motivations for The Independence View

One might be attracted by the Independence View if one does not believe in the existence of conceptual priority at all. On this radical view about conceptual relation, there does not exist any concept that is necessary for explaining a particular concept, and every concept can be understood independently of other concepts. It follows that even if two concepts are connected, neither is prior to the other.

Another motivation for accepting the Independence View is more modest. Those who are so motivated believe that we can master the concepts of moral wrongness and the appropriateness of having reactive attitudes independently and then come to realize that these concepts are related. In particular, we can understand an action to be morally wrong without thinking of the appropriate reactive attitude to hold as a response to the agent. Meanwhile, we can grasp the appropriateness of having reactive attitudes without invoking the deontic status of the relevant action.

2.3 Problem of the Independence View

Both claims (1) and (4) are true, but claim (5) is false. Recall that this claim (1) is true because the Non-Naturalist Reactive Attitude Approach is correct. Claim (4) states that the moral wrongness of A is not in virtue of the appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A. Since this claim is the negation of claim (2), which claim is false, it is true.

The problem of the Independence View is claim (5), which states that the appropriateness (permissibility, or obligation) for reacting people to feel the blame-sentiment (or express blame) towards S regarding S's performing A is not in virtue of the moral wrongness of A. I will provide two reasons for rejecting this claim.

First, claim (5) does not fit the way we normally talk. It is natural for us to say to another that "You deserve blame because you did the wrong thing," "You are to blame because you acted wrongly," and "She is right to resent you because you hurt her badly." These expressions all suggest that rightness/wrongness is conceptually connected to the

appropriateness of *expressing* blame, and rightness/wrongness explains why expressing blame is appropriate. Of course, this suggestion does not imply that *moral wrongness* is conceptually connected with the appropriateness of expressing the blame-sentiment, nor does it imply that *moral wrongness* is conceptually connected with the appropriateness of *holding* the blame-sentiment. However, it seems that we can infer the conceptual connection between moral wrongness and the appropriateness of holding the blame-sentiment in two steps.

We can first infer that *moral wrongness* and the appropriateness of expressing the blame-sentiment are conceptually connected. The primary reason is that, among all the reasons wrongness bears upon, moral reasons are especially weighty. It is reasonable to think that moral reasons outweigh all non-moral reasons in normal cases (e.g., when morality does not require one to make significant sacrifices). Accordingly, if an action is wrong all-things-considered, it is morally wrong in normal cases. Hence, if wrongness conceptually connects with the appropriateness of expressing the blame-sentiment, *moral wrongness* also conceptually connects with it in normal cases.

We can further infer that moral wrongness and the appropriateness of *holding* the blame-sentiment are conceptually connected. If an outward expression of a blame-sentiment is appropriate in a situation, then an inward feeling of the sentiment, as a weaker response to the action at stake, is also appropriate. The idea is that feeling a sentiment generates milder consequences than expressing the sentiment, and the appropriateness of expressing a sentiment involves more factors than the appropriateness of feeling a sentiment does. If I am right, since moral wrongness and the appropriateness of expressing the blame-sentiment are conceptually connected in normal cases, moral

wrongness and the appropriateness of *holding* the blame-sentiment are also conceptually connected in normal cases.

In a similar fashion, since expressions such as “You deserve blame because you acted wrongly” also suggest that we understand the appropriateness of expressing the blame-sentiment by understanding wrongness, we can infer that we understand the appropriateness of holding the blame-sentiment by understanding moral wrongness. This implies that the appropriateness of holding the blame-sentiment is conceptually grounded by moral wrongness. Hence, our way of speaking implies the negation of claim (5).

The second reason for rejecting claim (5) comes from my arguments in chapters 3 and 4. In Chapter 3, I argued that the appropriateness of feeling the blame-sentiment is not in virtue of the consequences of the action, the agent’s mental state (beliefs or intentions), the agent’s evidence, the agent’s available actions, and the agent’s character exhibited in performing the action, or the combination of two or more of the above. It follows that if the appropriateness of feeling the blame-sentiment is not in virtue of moral wrongness, it has to be in virtue of the options other than the ones I canvassed in Chapter 3. However, I do not think that there are any such options. In Chapter 4, I argued that Fitting Attitude Theories of Wrongness all provide circular explanation of the appropriateness of feeling the blame-sentiment. This problem indicates that moral wrongness is necessary for understanding such appropriateness, which entails the falsity of claim (5).

In summary, because claim (5) does not fit our way of speaking, and it is not successful to understand the appropriateness of feeling the blame-sentiment apart from moral wrongness, claim (5) is false.

I am aware of some counterarguments made by the proponents of claim (5). In particular, some defend claim (5) by arguing that rejecting it commits one to accepting certain unacceptable implications. A purported implication is that the notion of moral wrongness becomes mysterious.⁹² The idea is that we have no idea how to understand moral wrongness without understanding blameworthiness. Another purported implication is that we will be committed to strange things such as non-natural properties.⁹³ The idea is that if we insist that moral wrongness is conceptually prior to blameworthiness, then we will have to understand moral wrongness by recognizing the property of wrongness, and in this case the property is non-natural, which is absurd.

I will reply to these objections in Chapter 6, where I defend the Standard View. The reason is that the Standard View is under attack for various reasons, and addressing all the objections together is the main task of Chapter 6. For now, please allow me to postpone my replies to the proponents of claim (5). So far, I have rejected claim (5) mainly because my arguments in chapters 3 and 4. Hence, I claim that the Independence View is false.

3. Conclusion

In this chapter, I have argued against the Interdependence View and the Independence View. According to the Interdependence View, moral wrongness is defined in terms of the normative status of our having (or expressing) core negative reactive attitudes, and such normative status is also defined in terms of moral wrongness. I argue that this view

⁹² Skorupski, 2010b.

⁹³ Gibbard, 2008a.

is false because moral wrongness cannot be understood in terms of such normative status. According to the Independence View, moral wrongness is not defined in terms of the normative status of our having (or expressing) core negative reactive attitudes, nor is such normative status defined in terms of moral wrongness. I argue that this view is false because such normative status has to be understood in terms of moral wrongness.

This chapter will facilitate my defense of the Standard View in the next chapter. The Standard View holds that the normative status of our having/expressing core negative reactive attitudes is defined, at least partly, in terms of moral wrongness. Chapters 3, 4, and 5 amount to an argument for the Standard View because it is the only possible view left, after accepting the Non-Naturalist Reactive Attitude Approach and rejecting the Reactive Attitude View, the Interdependence View, and the Independence View.

Ch. 6 The Standard View

In this chapter, I will defend the Standard View. According to this view, moral wrongness, at least partly, defines the normative status (e.g., appropriateness) of our having (or expressing) core reactive attitudes (i.e., blame, resentment, indignation, and guilt), but not vice versa. I will first introduce this view, motivate it, and then defend it.

1. The Standard View

According to the Standard View, the normative status of our having or expressing core reactive attitudes is analyzed in terms of moral wrongness. It implies that moral wrongness is *conceptually prior* to the normative status of our having or expressing core reactive attitudes, but not vice versa. That is, if it is appropriate (permissible, or obligatory) for a person to hold/express an attitude towards an agent with respect to the agent's action, then the appropriateness (permissibility, or obligation) is explained, at least partly, by the fact that she inexcusably acts wrongly. Moreover, we need to understand the appropriateness (permissibility, or obligation) of holding/expressing an attitude in terms of the deontic status of the action at stake. However, we cannot and should not explain moral wrongness in terms of the appropriateness (permissibility, or obligation).

Before I discuss the Standard View in detail, I would like to distinguish conceptual priority from metaphysical (or ontological) priority, for they are closely related but easy to be confused. Conceptual priority is definitional priority, so it is about explaining and understanding concepts. For example, if belief is conceptually prior to knowledge, then we can define knowledge in terms of beliefs, but not the other way

around. That is, we can both explain the concept of knowledge in terms of the concept of belief and understand the concept of knowledge in terms of the concept of belief, but not vice versa. In other words, we understand the concept of knowledge because we understand the concept of belief, whereas we understand the concept of belief independently of our understanding the concept of knowledge. In contrast, metaphysical priority is property priority, so it is about explaining and understanding metaphysical relations.⁹⁴ For example, if physical properties are metaphysically prior to mental properties, then the existence of mental properties is explained in terms of the existence of physical properties, or mental properties are grounded in physical properties, but not vice versa. That is, physical properties are more basic than mental properties such that mental properties exist because physical properties exist, but physical properties exist independently of the existence of mental properties.

Conceptual and metaphysical priority are sometimes connected. For instance, it seems that belief is both conceptually and metaphysically prior to knowledge, for knowledge is a subset of belief and knowledge about something must be a belief about that thing in the first place. However, conceptual and metaphysical priority are not always connected. On the one hand, conceptual priority does not imply metaphysical priority. For instance, although child is conceptually prior to toy, the property of being a toy is not grounded in the property of being a child. On the other hand, metaphysical priority does not imply conceptual priority. For instance, physical properties might be metaphysically prior to mental properties, but mental properties need not be understood in terms of physical properties (e.g., we don't have to understand pain in terms of

⁹⁴ For a comprehensive introduction to metaphysical priority, see Schaffer, 2009.

particular brain states). Hence, since the Standard View merely claims that moral wrongness is *conceptually* prior to the appropriateness (permissibility, or obligation) of holding/expressing core negative reactive attitudes, it is silent on whether moral wrongness is *metaphysically* prior to such appropriateness (permissibility, or obligation). I focus on the issue of conceptual instead of metaphysical priority because theorists endorsing the Reactive Attitude View are interested in conceptual priority, and the issue of metaphysical priority will be a further project to work on.

The Standard View generally has the following form:

Standard View: It is appropriate (permissible, or obligatory) for reacting people to feel (or express) the blame-sentiment towards an agent, S, regarding S's performing an action, A, if and only if, and in virtue of the fact that, it was morally wrong for S to perform A, and S had no excuse.

Proponents of the Standard View include John Rawls (1971), arguably W. D. Ross (1930, 1939) and Immanuel Kant (1785, 1788).

Kant seems to adopt the Standard View when he introduces the moral feeling in the *Critique of Practical Reason*:

Moral feeling is the susceptibility to feel pleasure or displeasure merely from being aware that our actions are consistent with or contrary to the law of duty. Every determination of choice proceeds *from* the representation of a possible action *to* the deed through the feeling of pleasure or displeasure, taking an interest

in the action or its effect. The state of *feeling* here (the way in which inner sense is affected) is either *pathological* or *moral*. – The former is that feeling which precedes the representation of the law; the latter, that which can only follow upon it.⁹⁵

If we regard the moral feeling as a kind of reactive attitude, then Kant seems to hold that certain reactive attitudes (e.g., praise, blame) regarding some actions can only arise as a result of being aware of the deontic status of the actions. If so, then we can talk about the appropriateness or correctness of the moral feeling, for it is independent of anyone's mental state whether the relevant action is (objectively) morally right/wrong. In other words, we can understand the appropriateness of having certain reactive attitudes in terms of the rightness or wrongness of the relevant action.

Ross (1930, 1939) also seems to take the Standard View. In *Foundations of Ethics*, he says,

[S]orrow is *right* when one contemplates the death of a friend, and *wrong* when one contemplates the success of a rival. [...] We can still call grief the *right* or *fitting* emotion in certain situations [...] Its *fittingness* depends solely on the nature of the circumstances and not at all on his capacity or incapacity.⁹⁶

⁹⁵ Kant, 1788, 6:399, emphasis original.

⁹⁶ Ross, 1939, pp. 52-55, my emphasis.

On Ross's view, whether an emotion is fitting to have depends on the circumstances, namely, the relevant event involving personal relationship. Since he holds that moral rightness/wrongness is "a form of suitability" for the circumstances, the fittingness of an emotion seems to depend at least partially on moral rightness/wrongness.⁹⁷ If the emotions Ross discusses are reactive attitudes, then he seems to agree that the fittingness of having a reactive attitude can be defined in terms of moral rightness/wrongness.

Rawls (1971) is more explicit than Kant and Ross about the connection between reactive attitudes and wrongness. For Rawls,

[I]t is a necessary feature of moral feelings, and part of what distinguishes them from the natural attitudes, that the person's explanation of his experience invokes a *moral* concept and its associated principles. His account of his feeling makes reference to an acknowledged *right* or *wrong*.⁹⁸

Since moral feelings for Rawls are reactive attitudes, he takes reactive attitudes to presuppose moral rightness/wrongness. It follows that holding/expressing an attitude can be appropriate or inappropriate because there is an objective fact about whether the target agent acts rightly/wrongly.

In fact, the Standard View is compatible with both (the ontological sense of) moral realism and antirealism. Moral realism holds that there exist mind-independent

⁹⁷ Ross, 1939, p. 55.

⁹⁸ Rawls, 1971, p. 481, my emphasis.

moral facts.⁹⁹ It is natural, although not necessary, for a moral realist to believe that the moral wrongness of an action is both *metaphysically* and *conceptually prior* to the appropriateness (permissibility, or obligation) of holding/expressing blame, resentment, indignation, guilt towards the agent with respect of the action. Among realist theories, Cornell realists' (e.g., Nicholas Sturgeon 1985; Richard Boyd, 1988; David Brink 1989), Peter Railton's (1986), Russ Shafer-Landau's (2003), and David Enoch's (2011) theories all fit the Standard View.

Nonetheless, a moral realist need not accept the Standard View, since the Reactive Attitude View is also compatible with moral realism. For example, a moral realist may hold that the property of moral wrongness depends upon the property of blameworthiness, which in turn is mind-independent. If so, then she can coherently hold that the concept of moral wrongness depends upon the concept of blameworthiness. The point is that moral realism fits the Standard View well, although it does not require this view.

Not only does moral realism fit the Standard View, moral anti-realism fits this view as well. This is because the Standard View does not require the endorsement of mind-independent moral facts; rather, it merely requires the endorsement of truth-apt moral statements. As long as an anti-realist believes that moral statements of right and wrong have truth-value, then her theory is amenable to the Standard View. For instance, according to John Mackie's error theory (1977), there are no mind-independent moral facts, but moral statements have truth-value, whose obtaining requires such moral facts.

⁹⁹ For the four senses (semantic, ontological, metaphysical, and normative) of moral realism, see Finlay, 2007.

Hence, “positive” moral statements (e.g., the statement that certain action is morally right/wrong) are always false for Mackie. He can coherently hold that moral wrongness is conceptually prior to the fittingness of holding/expressing reactive attitudes.

The Standard View can be either cognitivist (i.e., moral statements express propositions) or non-cognitivist. For example, a norm-expressivist can hold that the norm of moral wrongness is prior to the norm of the fittingness of having reactive attitudes. Hence, which is the appropriate norm to adopt concerning reactive attitudes is a matter of which attitude fits the norm of moral wrongness. I will focus on cognitivism for simplicity.

In the last chapter, I argued that it is more plausible for the Reactive Attitude View to appeal to *having* reactive attitudes than *expressing* reactive attitudes, and it is more plausible for it to appeal to the *appropriateness* of having reactive attitudes than the *obligation* or *permissibility* of having reactive attitudes. This also holds for the Standard View for the same reason, namely, (1) *having* an attitude is mainly backward-looking while *expressing* an attitude can be mainly forward-looking, and (2) the *appropriateness* of having an attitude is merely backward-looking while the *obligation* or *permissibility* of having an attitude is both forward-looking and backward-looking. Hence, in the remainder of this chapter, I will only discuss the versions that discuss the *appropriateness* of *having* reactive attitudes.

In the following section, I will motivate the Standard View.

2. Motivations for the Standard View

The Standard View is attractive for several reasons. First, the notion of moral wrongness seems intelligible without any appeal to the fittingness of holding reactive attitudes towards one another. It is true that we may need the practice of feeling resentment or indignation to deepen our understanding of moral wrongness, but the practice does not seem necessary for us to gain the concept of moral wrongness. In any case, we normally have the cognitive capacity to understand abstract notions of moral obligation, moral permissibility, and moral prohibition. Along this line, some moral intuitionists (e.g., Ross, 1930, p. 29; Prichard, 1937) even go as far as to claim that the notion of rightness is self-evident such that we can just use intuition to grasp whether a given action is morally right.¹⁰⁰

Second, we have seen that the Reactive Attitude View encounters insurmountable difficulties in specifying when it is appropriate to hold core negative reactive attitudes, namely, the wrong kind of reason problem. For example, Darwall's account fails to specify when it is appropriate to hold a blame-sentiment. In contrast, the Standard View offers a straightforward solution to this problem. That is, it is appropriate to blame someone for an action if and only if the relevant action is objectively morally wrong and the agent has no excuse. This solution is appealing in that it is simple, clear, and practical. If it works, then the Standard View is more plausible than the Reactive Attitude View in this respect.

¹⁰⁰ Strictly speaking, Ross thinks that we can directly know whether a *prima facie* duty (a duty in terms of a, not all, moral consideration) exists, but we cannot know for sure the absolute duty (a duty in terms of conclusive moral considerations) of a particular situation.

Third, if I am right that the Reactive Attitude View is circular in analyzing moral wrongness in terms of the normative status of having core negative reactive attitudes, then the Standard View has to be correct. This is because the Standard View analyzes such normative status precisely in terms of moral wrongness.

In Section 3, I will defend two arguments for the Standard View, an overarching argument and an argument based on the observation that wrongness connects with negative fitting reactive attitudes.

3. Argument for the Standard View

3.1. Argument from Elimination

My first argument is overarching. Its basic idea is that the Standard View has to be correct when the Non-Naturalist Reactive Attitude Approach is correct and all the other views taking this approach are false. Let's call this argument *the Argument from Elimination*.

Here is a formulation of this argument:

E1: If the Reactive Attitude View, Independence View, and Interdependence View are all false, then the Standard View is correct.

E2: The Reactive Attitude View, Independence View, and Interdependence View are all false.

C: The Standard View is correct.

E1 follows from the acceptance of the Non-Naturalist Reactive Attitude Approach. The Reactive Attitude View, Independence View, Interdependence View, and the Standard View exhaust all the space of this approach, so the Standard View is correct if all the others are false. E2 states that all the other views are indeed false. The conclusion, C, follows from E1 and E2.

The argument from Elimination is supported by chapters 3 through 5. I defended the Non-Naturalist Reactive Attitude Approach in Chapter 3. Hence, E1 is defended. I argued against the Reactive Attitude View in Chapter 4. And I argued against the Independence View and Interdependence View in Chapter 5. Hence, E2 is defended. Let me briefly explain why E2 is true.

The Reactive Attitude View holds that blameworthiness is conceptually prior to moral wrongness. It has to account for the thing blameworthiness is based upon. Possible candidates include several aspects of the action: the deontic status of the action, the moral value of the action, the agent's intentions of performing the action, and the agent's character exhibited in performing the action, or the combination of two or more of the above.

If blameworthiness is based on the deontic status of the relevant action, then the view that blameworthiness is conceptually prior to moral wrongness is circular, so this option is closed. If blameworthiness is based on the moral value of the action, then it avoids being circular, but the talk of blameworthiness in terms of moral value fails to capture the essential aspect of blameworthiness. Hence, the option of moral value is eliminated. Third, the appeal to intention is irrelevant to our notion of blameworthiness, i.e., objective blameworthiness, which is blameworthiness about the real situation. In

particular, blameworthiness in terms of intention concerns *subjective* blameworthiness, i.e., blameworthiness with respect of the situation the agent wants to bring about. Hence, the option of motive is eliminated. Lastly, it does not work to say that blameworthiness is based on the agent's character exhibited in performing the action because Civilian is blameworthy even if her character is just when she helps Criminal get out of the jail due to a mistaken belief that Criminal is innocent. Since all these options have problems, the combination of any two or more of them does not work, since that will inherit the problems.

The ways I consider cashing out blameworthiness are prominent ones. I admit that there might be other ways, and one of which might be successful, but I doubt that there are. It might help briefly look at Fitting Attitude Theorists' proposal I discussed in Chapter 4. First, Skorupski interprets blame as a tendency to withdraw, but does not clearly account for where blame fits, so his proposal is at best incomplete. Next, Gibbard regards blameworthiness as the warrant of guilt and resentment, but his account of warrant presupposes objective moral wrongness. Lastly, Darwall explains blameworthiness in terms of second-personal reasons, but such reasons implicitly appeal to moral wrongness. Thus, I conclude that it is very unclear how to understand blameworthiness apart from moral wrongness and make blameworthiness conceptually prior to moral wrongness.

Let's turn to the Interdependence View and Independence View. The Interdependence View holds that blameworthiness and moral wrongness are inter-defined. But since moral wrongness cannot be non-circularly defined in terms of blameworthiness, this view is false. As for the Independence View, it holds that neither

blameworthiness nor moral wrongness is defined in terms of the other. But since any definition of moral wrongness in terms of blameworthiness is circular, blameworthiness is defined in terms of moral wrongness. Thus, the Independence View is false.

If my arguments in chapter 3 through 5 succeed, then the Argument from Elimination succeed, and hence the Standard View is correct. The Argument from Elimination indirectly argues for the Standard View. I am aware that many objections are levelled against the Standard View, so my defense of the Standard View will not be satisfying if I do not address them. Hence, I will take up the objections in the next subsection, where I directly defend for the Standard View.

3.2 Argument from Connection

My second argument for the Standard View is an inference to the best explanation. Since it is fitting to blame an agent if and only if she acts wrongly with no excuse, the notions of blameworthiness and wrongness probably connect in certain ways. One explanation, provided by the Standard View, is that moral wrongness is conceptually prior to blameworthiness. I shall argue that this explanation best accounts for the connection between blameworthiness and moral wrongness.¹⁰¹ Call this argument *the Argument from Connection*.

¹⁰¹ For simplicity, I use blameworthiness as an example of the fittingness of having a core negative reactive attitude. Similar arguments can be made about the fittingness of holding other core negative reactive attitudes (e.g., resentment). The difference is that the reacting people are connected to the target agent in certain ways. For example, if the attitude we discuss is resentment, then the observation will roughly be that it is generally appropriate for a reacting person to feel resentful towards an agent who acts wrongly *and wrongs her* with respect of the action.

My argument runs as below:

P1: An agent is blameworthy if and only if she inexcusably acts wrongly.

P2: The Standard View provides a good explanation for P1.

C: The Standard View is correct.

Although the Argument from Connection is not valid, it offers strong reasons for the conclusion. In the rest of this subsection, I will defend this argument.

P1 describes the observation that an agent is blameworthy if and only if she inexcusably acts wrongly. It is a conceptual claim. ‘Blameworthiness’ here refers to the appropriateness of the sentiment of blame. P1 expresses the judgment that it is appropriate to express the sentiment of blame towards an agent who inexcusably acts wrongly. Because our concern is about the connection between moral wrongness and blameworthiness, we can qualify P1 in restricting wrongness to moral wrongness. As for blame, I am aware that philosophers are divided about whether it is characteristically moral. As I clarified in Chapter 1, I will focus on cases where blame is connected with moral wrongness, without taking a stance on whether blame can be characteristically non-moral. Qualified this way, P1 is a version of the Non-Naturalistic Reactive Attitude Thesis, namely, it is morally wrong for an agent, S, to perform an action, A, if and only if were S to have no excuse then it would be appropriate (permissible, or obligatory) for any reacting person to feel (or express) the blame-sentiment towards S regarding S’s performing A. I have defended this thesis in Chapter 3. Thus, I will assume that it is right here. P1 is not only accepted by theorists of the Standard View, but also theorists of the

Interdependent/Independent View and the Reactive Attitude View. Indeed, Skorupski takes the conceptual connection between moral wrongness and blameworthiness to be *a priori* (2010b, p. 293).

The crucial premise is P2, which states that the Standard View provides a good explanation for P1. According to the Standard View, moral wrongness is conceptually prior to blameworthiness. This explanation is good because it is informative, simple, and clear. First, this explanation informs us why the blame-sentiment is fitting if an agent inexcusably acts wrongly: namely, moral wrongness is conceptually prior to blameworthiness. When concept A is prior to concept B, it is not surprising that concept B is invoked every time concept A occurs in a certain fashion. Second, this explanation is simple and clear. In particular, all it tells us is that moral wrongness and blameworthiness are conceptually connected in a certain way. We seem to have no difficulty in understanding this idea (objection about how to understand moral wrongness will be addressed in the next section). Thus, I conclude P2 to be true.

Since both P1 and P2 are true, it provides good reason for the conclusion, namely, the Standard View is correct. Nevertheless, since the Argument from Connection is an inference to the best explanation, it does not firmly establish the conclusion.

Indeed, both the Reactive Attitude View and the Interdependence View can explain P1 to some extent. The Interdependence View holds that moral wrongness and blameworthiness are defined in terms of each other. According to this View, excusably morally wrong actions entail blameworthiness, and blameworthiness entails moral wrongness. So P1 is explained. According to the Reactive Attitude View, blameworthiness is conceptually prior to moral wrongness. It explains why an agent acts

wrongly if she is blameworthy and has no excuse. In fact, Skorupski argues for the Reactive Attitude View by appeal to P1.

I have rehashed my objections to the Reactive Attitude View and the Interdependence View in Section 3.1. Regarding the explanatory power for P1, the Standard View and the Reactive Attitude View are on a par. Specifically, the Standard View explains why an agent is blameworthy if she excusably acts wrongly, but does not explain why an agent acts wrongly if she is blameworthy and has no excuse. In contrast, the Reactive Attitude View explains the latter but not the former. In the following section, I will compare the merits and costs of the Standard View and the Reactive View. Meanwhile, I will make sense of the Standard View from various criticisms.

4. The Standard View vs. The Reactive Attitude View

The Standard View and the Reactive Attitude View are both appealing. Like the Interdependence View, they both explain the conceptual connection between moral wrongness and blameworthiness. Unlike the Interdependence View, they are not blatantly circular. Despite these merits, they both have drawbacks. In particular, the Reactive Attitude View is challenged to be circular while the Standard View is allegedly mysterious, among others. In the rest of this section, I will explore the options to meet these challenges. My claim is that neither challenges can be met. However, in light of the fact that being mysterious is preferable to being circular, the Standard View is preferable to the Reactive Attitude View.

4.1 The Reactive Attitude View Revisited

The Reactive Attitude View is challenged by two problems: the wrong kind of reason (WKR) problem, and circularity.

4.1.1 WKR Problem

Roughly, the WKR problem is that the Reactive Attitude View cannot distinguish the right kind of reason for having an attitude from the wrong kind. The most leading approach to this problem is by appeal to Parfit's distinction between object- and state-given reasons. As the names suggest, object-given reasons are reasons concerning the object, while state-given reasons are reasons concerning the state of having an attitude. If the two types of reasons are distinct, then presumably we can identify the right kind of reasons by looking for object-given reasons, and the wrong kind of reasons by looking for state-given reasons.

I argued in Chapter 4 that the object-/state-given reasons distinction is unclear. I also argued that some wrong kind of reasons do not seem state-given. Here, I'll focus on a systematic problem of the approach.

Schroeder (2012b, 2013) points out that if the approach is correct, then state-given reasons cannot be the right kind of reasons. However, sometimes a state-given reason is a reason for us to remain neutral—neither to have a particular attitude nor to have the opposite attitude. Since sometimes such a reason is the right kind of reason, the appeal to the object-/state-given reasons distinction fails. For example, a Catholic hears that a woman had an abortion. Because of his Catholic beliefs, he does not approve of abortion. However, he does not know why the abortion happened. If it happened because the

woman did not want to have the baby, he would feel the blame-sentiment towards her. But if it happened because the woman had cancer of the uterus, he would not feel so. Now, a TV station will talk about the details of this case. Suppose the Catholic can only know about the case by watching TV in a few hours. In this case, the information about why the woman had abortion is crucial, and the fact that the TV station will discuss it provides a reason for the Catholic to withhold judgments and attitudes. That reason is a state-given reason, since it is provided by situational factors. It is also the right kind of reason, since it bears upon a crucial aspect of the woman's abortion. Cases like this constitute counterexamples to the approach appealing to the object-/state-given reasons distinction.

The failure of the approach does not entail the failure of the Reactive Attitude View to solve the WKR problem. However, it does make the prospect of solving the problem dim, for it is probably the best-developed approach.

4.1.2 Circularity

Compared to the WKR problem, circularity is a more serious charge against the Reactive Attitude View. In Chapter 4, I argued that Darwall's, Skorupski's, and Gibbard's accounts of wrongness are all circular. Yet, I have not shown that every account of the Reactive Attitude View is circular. Hence, I need a general argument to show that. My strategy is to work out a plausible account of the Reactive Attitude View, and show that this account is circular.

Since the Reactive Attitude View analyzes moral wrongness in terms of blameworthiness, it needs to explain blameworthiness without the appeal to moral

wrongness. Nowadays, a number of theorists use the notion of reason to explain moral notions (Scanlon, 1998; Skorupski, 2010b). Hence, let's try to develop an account using reason as the primitive notion. The idea is to explain moral wrongness in terms of blameworthiness, which is in turn explained in terms of reasons. Here is one formulation: A person is blameworthy if and only if there is good reason to blame her. We can interpret 'good reason' in different ways. For example, for Scanlon, the existence of a good reason to blame somebody amounts to the fact that the person impairs her relationship with those who blame her because of her action. For our purpose, we can leave the details about good reasons open. If we go along with this notion of blameworthiness, we can understand moral wrongness in this way: a person acts morally wrongly if and only if, and in virtue of, were she to have to excuse, there would be good reason to blame her. Let's examine this account.

On the face of it, this account is not circular. This is because the definition of blameworthiness seems to have no appeal to moral wrongness. However, note that the definition of blameworthiness has to have a moral component.¹⁰² This is because the notion of blameworthiness is used to define moral wrongness, not wrongness in general. As we know, a person can be blameworthy for non-moral reasons (such as failing to follow one's schedule to get up at 7AM, which normally has no connection with morality). Accordingly, we have to qualify the reasons for blaming to be the moral reasons. But once we do that, we run the risk of being circular. In particular, we have to distinguish moral reasons from non-moral reasons, which allows us to distinguish the blame that connects with moral wrongness. The question is: How do we distinguish

¹⁰² Thanks to Robert Johnson for this suggestion.

moral from non-moral reasons? If we simply distinguish all the moral considerations from non-moral ones, then what we do amounts to distinguish all the considerations for moral evaluations. Since moral wrongness is a moral evaluation, this method implicitly appeals to moral wrongness.

One might argue that we can only be blameworthy for moral reasons. If this line is right, then there is no need to distinguish moral from non-moral reasons. Yet this line implies that every situation is a moral one, which is implausible. We blame others for failing moral standards, religious standards, institutional standards, personal standards, etc. It is not plausible that all these standards reduce to moral ones. Hence, I do not think that this line succeeds.

Let's take a look of why this trial account of the Reactive Attitude View fails. An account of the view has to explain blameworthiness without the appeal to moral wrongness. But in light of the fact that one can be blameworthy for moral or non-moral reasons, this account has to distinguish moral reasons from non-moral reasons. That distinction cannot be made without the appeal to moral wrongness, so the trial account is circular. Actually, any account of the Reactive Attitude View runs into the same problem. Regardless of which notion is rendered basic, it has to distinguish moral blameworthiness from non-moral blameworthiness, which cannot be done without the help of moral wrongness. Thus, the Reactive Attitude View cannot avoid being circular.

What about the Standard View? Does it fare better as compared to the Reactive Attitude View?

4.2 The Standard View

The Standard View is challenged from three concerns: the ability to answer the open question argument, the ability to accommodate folk psychology, and the ability to be clear.

4.2.1 Open Question Argument

The first concern comes from Moore's open question argument. A skeptic of the Standard View might say: "If the Standard View is correct, then an agent is blameworthy whenever she inexcusably acts wrongly. However, if an agent inexcusably acts wrongly, it feels like an open question whether she is blameworthy. For example, if an agent harms me without any excuse, we can still sensibly ask 'Is she blameworthy?'" A more radical skeptic (an extremely nice Christian, for example) might even insist that she is not blameworthy at all.

If this objection works, then it reveals that blameworthiness cannot be analyzed in terms of moral wrongness and the absence of excuse, which entails the falsity of the Standard View. However, I don't think that the objection works. First, Moore's original open question argument is supposed to reveal a conceptual gap between non-natural notions and natural notions.¹⁰³ For example, it seems to always make sense to ask "This thing is desired by most people, but is it good?" The feeling of this conceptual gap is partially caused by the difference we feel between the non-natural concept of goodness and the natural concept of being-desired. In contrast, the Standard View analyzes a non-natural notion, blameworthiness, in terms of another non-natural notion, moral wrongness.

¹⁰³ Thanks for Matt McGrath for this suggestion.

Since these two notions are both non-natural, we cannot feel any conceptual gap in terms of the natural/non-natural distinction. Thus, the force of the open question argument weakens when the argument applies to the Standard View.

Second, the open question argument after all does not conclusively disprove the position at stake. The reason is that this argument concerns our cognitive capacity of recognizing conceptual/metaphysical connections and this capacity is limited. For instance, if a person does not know that water is composed of H₂O, then she may think that it is an open question to ask that “This is H₂O, but is it water?” She will know that the question is not open until she learns that water consists of water. In the same fashion, if a person does not know that one plus one equals two, then she may think that it is an open question whether one plus one does equal two. And she will know that the question is not open either, when she learns further what two actually means. Likewise, since multiple questions about moral wrongness and blame are in dispute, we don’t know everything about moral wrongness and blameworthiness. Hence, if a skeptic of the Standard View thinks that it is an open question whether an agent inexcusably acts wrongly is blameworthy, it does not prove that the Standard View is false. It is plausible that the question is not really open, although the skeptic feels that it is.

4.2.2 Folk Psychology

The second objection I will address comes from folk psychology. If the Standard View is correct, assuming that our folk psychology is not vastly mistaken, it should generally fit the way we talk about wrongness and blameworthiness. Some talk of this kind is consistent with the Standard View. For example, it is natural to say that “You are

blameworthy because you did something wrong,” or “You are blameworthy because you wronged/hurt/harmed her.” These expressions imply that understanding blameworthiness is based upon understanding wrongness.

However, some skeptics argue that we need the concept of blame, guilt, etc. to acquire the concept of moral wrongness. If they are right, then moral wrongness cannot be conceptually prior to blameworthiness (or the fittingness of core negative reactive attitudes), which entails the failure of the Standard View. Skorupski puts the objection this way:

[...] we can elucidate a concept by saying what is required for possession of the concept, rather than by the semantic route of explicit definition within the existing resources of a language. What distinguishes moral wrongness as a *normative* concept is that you have the concept only if you are sensitive to reasons for the blame sentiment.¹⁰⁴

Skorupski correctly points out that we have to be sensitive to reasons for the blame-sentiment as reasons for blame. But he is wrong in holding that we have to possess the concept of blameworthiness to master the concept of moral wrongness. First, being sensitive to reasons for the blame-sentiment does not imply possessing the concept of blame. Second, empirical evidence suggests that children acquire the concept of moral wrongness prior to acquiring the concept of guilt. Let me elaborate.

¹⁰⁴ Skorupski, 2010b, p. 293, emphasis original.

Just because we are sensitive to reasons for the blame-sentiment, it does not mean that we have already possessed the concept of blame. By analogy, just because a child is sensitive to the reasons that an apple is tasty, it does not mean that she has possessed the concept of tasty. For instance, a 3-year-old girl may be inclined to choose a Gala Apple that is redder than the other (because she likes red objects), so she is sensitive to the reasons for a Gala Apple to be tasty. However, she need not understand what tasty means, for that is a separate issue from being sensitive to tasty-bearing reasons. Likewise, what really follows from our sensitivity to reasons for blame is merely that we have the cognitive capacity to distinguish the reasons bearing on blame from those that do not.

Furthermore, empirical evidence counts against the claim that we need the concept of blame, guilt, etc. to acquire the concept of moral wrongness. Studies suggest that children acquire the concept of moral wrongness between the ages of three and three and a half, but they acquire the concept of guilt between four and eight.¹⁰⁵ Moreover, autistic individuals can distinguish moral violations from conventional ones, which suggest that they have the concept of moral wrongness. But they do not understand pride, a complex emotion that seem to emerge at about the same time as guilt, which suggests that they do not have the concept of guilt. Thus, it does not seem that we need to possess the concept of blame, guilt, etc. to possess the concept of moral wrongness.

Two kinds of responses are possible. First, Skorupski might argue that children between the ages of three and three a half only have the capacity to apply the term ‘wrong,’ but they have not genuinely possessed the concept of moral wrongness.¹⁰⁶

¹⁰⁵ Nichols, 2004, p. 90.

¹⁰⁶ Thanks to Robert Johnson for this suggestion.

This response is problematic. According to an influential school in current developmental science, the core knowledge of folk physics and psychology emerges early. The core knowledge might be enriched through development, but it is deemed to remain unrevised into adulthood and to continue to guide judgments for adults.¹⁰⁷ Evidence suggests core moral judgment is stable across cultures. If the thought of this school is correct, then the children between three and three and a half do have moral knowledge about rightness and wrongness, which implies that they have the concept of moral wrongness.

Furthermore, it is risky for theorists of the Reactive Attitude View to put a high bar over when a person acquires the concept of moral wrongness and makes genuine moral judgments. In particular, evidence suggests that sophisticated moral judgment is not stable across cultures. Specifically, there is a wide spectrum over how adults reason in different cultures. In some cultures, adults reason in terms of reciprocal needs, whereas in other cultures, adults reason in terms of individual rights or universal moral principles.¹⁰⁸ The Reactive Attitude View is supposed to apply across cultures, so it does the view no good by making the possession of moral concepts a sophisticated process.

Another response is to argue that although children having the concept of moral wrongness do not have the concept of guilt, they have certain emotion concept that are morally laden. And such an emotion concept helps children acquire the concept of moral wrongness. Gibbard (2006c) suggests such a picture:

¹⁰⁷ Nichols, 2004, p. 93.

¹⁰⁸ Nichols, 2004, pp. 93-94.

We can share with a wide range of people, then, a set of pared-down narrowly moral concepts. I'll call these "near-moral" concepts. An action is outrageous, in this near-moral sense, just in case outrage over it is warranted on the part of others. Suppose, then, as may well be, that the concept of outrage is a human universal, and that the concept emerges fairly early in a child's development. Then, for all we have seen, the quasi-moral concept of being outrageous may be available to all adults and to all children at a fairly early stage. All they need is a concept of reprehension or outrage and a concept of warrant, and they have the ingredients for a near-moral concept of being reprehensible.¹⁰⁹

Gibbard suggests that a child could acquire the concept of moral wrongness through the possession of the concepts of outrage and warrant, while possessing the concept of outrage does not require possessing the concept of guilt. In his view, it is quite plausible that we acquire the concept of moral wrongness in an intuitive way first (through our feelings of outrage), and then we come to understand what seems wrong can be distinct from it really being wrong.

Gibbard's picture sounds plausible, but cannot be proved without empirical evidence. I agree that we have to understand some basic feelings (e.g., pain) to understand moral wrongs, but I am skeptical that we have to have the concept of outrage before we have the concept of moral wrongness. To me, we can connect moral wrongness with various feelings. So, complex feelings like outrage and guilt just do not seem necessary for us to understand moral wrongness.

¹⁰⁹ Gibbard, 2006c, p. 203.

Thus, I conclude that the evidence from folk psychology does not count against the Standard View.

4.2.3 Mysteriousness

The last objection is that the Standard View makes the notion of moral wrongness mysterious or obscure. This objection is the most serious one because the Standard View does not work if there is no intelligible way to understand moral wrongness apart from the fittingness of core negative reactive attitudes. Let's take a look at a couple of instances of this objection. Skorupski argues that the contemporary philosophical conception of moral obligation and wrongness are problematic. In particular, he says,

[...] moral obligation has become a problematic concept, at least in contemporary philosophical discussion and perhaps beyond. I do not mean that the experience of moral obligation is unfamiliar to most people. I mean that there is uncertainty about what to make of it. This uncertainty extends to the whole distinctive circle of moral concepts: moral wrongness; permissibility; obligation...¹¹⁰

Similarly, Gibbard states,

An answer to the question of how to live, they [many philosophers] would say, just is a belief as to what we ought to do and what it's okay to do. Now of course anyone who says this has the burden of explaining what 'ought' and 'okay' mean.

¹¹⁰ Skorupski, 2010b, p. 290.

If they can't, or if their answer involves strange and incredible things like non-natural properties, I then say that my direction of explanation is better.¹¹¹

The central idea of the objection is that concepts such as moral wrongness and permissibility are so abstract that they, taken alone, are intractable or empty. To people like Skorupski and Gibbard, moral wrongness is meaningless without the appeal to blameworthiness.

The concern that the notion of moral wrongness seems mysterious can be interpreted in two ways. The first interpretation is that the notion of moral wrongness is abstract. Strictly speaking, moral wrongness means moral impermissibility, a sense of 'should not.' For example, when we say that "You acted wrongly," we intend to express that "You shouldn't have done that" or "You ought not to have done that." But how to make sense of that? Can we analyze moral wrongness in any non-trivial way?

Two positions are promising in my view. Both views preserve some mysteriousness in the notion of moral wrongness, but they are intelligible. The first position is Ross's. On this position, the notion of moral wrongness is just undefinable (except in narrowly deontic terms, such as not permissible). This notion is thin and primitive. Our conceptual framework has to start from some primitive notions, and moral wrongness is one of them. We have numerous primitive and abstract notions, such as number and reason. Although they are mysterious, we understand them and use them in our daily life. Likewise, moral wrongness could be mysterious and intelligible at the same time.

¹¹¹ Gibbard, 2008a, p. 20.

The second position is that the notion of moral wrongness is definable normatively, other than the normative status of holding/expressing core negative reactive attitudes. The idea is that something other than such normative status is conceptually prior to moral wrongness. Perhaps moral agency plays this role. Michael Smith proposes a framework starting from the notion of ideal moral agents. If one could flesh out moral wrongness using such notion, moral wrongness could be nontrivially analyzable.

In summary, neither position explains away the mystery of the notion of moral wrongness, but they can at least somehow make sense of it. I am more inclined to take the first position, but I need not defend it here.

The second interpretation of the objection is that moral wrongness does not seem to refer to any natural properties. Mackie's argument from queerness makes us aware of the particularity of moral properties. This concern is well raised, but it can be answered.

First of all, holders of the Standard View can construe moral properties as natural properties, which makes the view immune from the challenge of obscurity. For instance, Railton (2003) construes moral properties as what he dubbed 'social rationality,' namely, "what would be rationally approved of were the interests of all potentially affected individuals counted equally under circumstances of full and vivid information."¹¹² Roughly speaking, he thinks that what people rationally agree upon in a society is right, what people rationally reject in a society is wrong, and what people rationally disagree about in a society is indeterminate in terms of the deontic status. Here, this notion of wrongness is not obscure because, given appropriate specification, people's rational judgments are intelligible.

¹¹² Railton, 2003, p. 22.

Second, even if holders of the Standard View construe moral properties as non-natural, the view can still be intelligible. For instance, Shafer-Landau claims that moral properties are *realized* by “a particular concatenation of descriptive facts,”¹¹³ even if they cannot be so defined. The realizability claim makes moral wrongness intelligible because natural properties ground or determine wrongness. In other words, without the appeal to the normative status of holding reactive attitudes, wrongness is intelligible by the appeal to the relation to natural properties.

In summary, the Standard View withstand the objections from the open question argument and folk psychology, but it is subject to being mysterious. As we have seen, the Reactive Attitude View might be able to withstand the WKR problem, but it cannot avoid being circular. Between the Standard View and the Reactive Attitude View, the former is preferable because being mysterious is a lighter burden than being circular. Thus, I conclude the Standard View to be correct.

5. Conclusion

In this chapter, I have defended the Standard View, which holds that moral wrongness is conceptually prior to the appropriateness (permissibility, or obligation) of holding/expressing core negative reactive attitudes. My main argument is that the Standard View provides a good explanation for the observation that an agent is blameworthy if and only if she inexcusably acts wrongly. The Standard View has been challenged from the open question argument and folk psychology, but these challenges can be successfully met. The Standard View remains mysterious, but it is a lesser evil

¹¹³ Shafer-Landau, 2003, p. 77.

compared to being circular, which is a problem of the Reactive Attitude View. Hence, the Standard View is better than the Reactive Attitude View. In the next chapter, I will draw the implications of the Standard View.

Ch. 7 Conclusion

I have argued for the Standard View, according to which the fittingness of holding core negative reactive attitudes (i.e., resentment, indignation, guilt, and blame) is at least partly defined in terms of moral wrongness, but not vice versa. In this chapter, I will talk about the implications of the Standard View, especially concerning moral responsibility.

1. The Main Results

Before discuss the implications, I will first recapitulate the big picture. The main question of my dissertation is the relationship between moral wrongness and reactive attitudes. I have argued that they are connected conceptually, not empirically. That is, our *empirical* disposition to hold/express negative reactive attitudes under certain circumstances is *not* always connected with an agent's action being morally wrong. But moral wrongness is *conceptually* connected to the appropriateness or permissibility of holding/expressing a reactive attitude. Several views can explain this conceptual connection, among which the Reactive Attitude View and the Standard View do the best job. The Reactive Attitude View holds that moral wrongness is defined, at least partly, by the normative status of our having (or expressing) core negative reactive attitudes, but not vice-versa. In contrast, the Standard View holds that such normative status is defined, at least partly, by moral wrongness. When we compare the strengths and weaknesses of these views, we find that the Standard View is better. Hence, I conclude that the Standard View is correct.

2. Reactive Attitudes and Moral Permissibility, Moral Obligation, etc.

The most straightforward implications of the Standard View are the conceptual connection between fitting core negative reactive attitudes and moral notions such as moral permissibility and moral wrongness. We can derive such conceptual connection because the Standard View bears on the conceptual connection between moral wrongness and fitting core negative reactive attitudes, while moral wrongness, moral permissibility, and moral obligation etc. are conceptually connected.

For instance, since moral permissibility just is the absence of moral wrongness, we can simply replace ‘morally wrong’ in the formulation of the Standard View with ‘not morally permissible’ and infer that it is appropriate for reacting people to feel the blame-sentiment towards an agent, S, regarding S’s performing an action, A, if and only if, and in virtue of the fact that, it was not morally permissible for S to perform A, and S had no excuse. Thus,

Moral Permissibility: It is *morally permissible* for S to perform A if and only if it would not be appropriate for reacting people to feel the blame-sentiment towards S regarding S’s performing A.

Likewise, we can derive an account of conclusive moral obligation. In particular, S has a *conclusive moral obligation* to perform A, if and only if, it is morally permissible for S to perform A and it is morally impermissible for S not to perform A. Thus, we arrive at this:

Conclusive Moral Obligation: S has a *conclusive moral obligation* to perform A, if and only if (1) it would not be appropriate for reacting people to feel the blame-sentiment towards S regarding S's performing A, and (2) were S to have no excuse, it would be appropriate for reacting people to feel the blame-sentiment towards S regarding S's failure to perform A.

Since moral optionality is moral permissibility without moral obligation, we can derive that:

Moral Optionality: It is *morally optional* for S to perform A, if and only if (1) it would not be appropriate for reacting people to feel the blame-sentiment towards S regarding S's performing A, and (2) it would not be appropriate for reacting people to feel the blame-sentiment towards S regarding S's failure to perform A.

Since moral supererogation is moral optionality plus moral value, we can arrive at an account of moral supererogation.

In the next section, I will talk about the implications concerning moral responsibility.

3. Moral Responsibility

While some philosophers continue to regard moral responsibility to be a generic notion, more draw a distinction between two senses of moral responsibility: attributive responsibility (or agent responsibility) and substantive responsibility (or

accountability).¹¹⁴ A person is attributively responsible for an action if the action can be attributed to her as an agent, that is, the action reflects the agent's mental states. In contrast, a person is accountable for an action if she should bear certain consequences (reward or punishment) as a result of her action. Unlike attributive responsibility, accountability is essentially a social notion because it is about the consequences one should bear as a member of the society.

Among those distinguishing the two senses of moral responsibility, some (Scanlon, 1998) regard praise and blame to be merely concerned with attributive responsibility, while others (Watson, 1996) hold praise and blame to be concerned with both attributability and accountability. In the following, I will take the latter view to be correct. The main reason is that our judgments about blame are pulled in two ways. On the one hand, we seem to blame others only if their actions are beyond rational control, roughly speaking. This phenomenon seems to imply that blame is connected with accountability, since one is not accountable for her action if her action is beyond rational control. On the other hand, we sometimes blame others even if their actions are beyond rational control. This phenomenon seems to imply that blame is connected with attributability, since whether an action can be attributed to an agent does not depend upon whether she can rationally control the action. I believe that both phenomena are natural and sensible, so I believe that blame can be connected with either attributability or accountability.

¹¹⁴ Watson, 1996. Scanlon, 1998. Eshleman, 2009. Shoemaker thinks that moral responsibility has three senses. See Shoemaker, 2007.

Attributive responsibility is connected more with *having* reactive attitudes than with *expressing* attitudes. This is because the attitudes are held towards the agent in question in the sense that the action in question can be attributed to the agent. Granted, we might express the attitude to someone attributively responsible for an action, but we do not normally do so. The reason is that certain attitudes, especially blame, constitute sanction, while the objects of sanction are not normally those who cannot rationally control their actions. In contrast, accountability is connected more with *expressing* reactive attitudes than *having* attitudes. This is because the attitudes are expressed towards the agent in question in the sense that the action in question makes the agent an appropriate candidate for bearing the attitudes. Expressing negative attitudes, compared with punishment, is a weak kind of sanction.

This is the formulation of the Standard View:

Standard View: It is appropriate (permissible, or obligatory) for reacting people to feel (or express) the blame-sentiment towards an agent, S, regarding S's performing an action, A, if and only if, and in virtue of the fact that, it was morally wrong for S to perform A, and S had no excuse.

It seems that the version talking about having or feeling an attitude is about attributive responsibility, while the version talking about expressing an attitude is about accountability. However, since the Standard View considers excuses, whereas an action being beyond rational control is an excuse, it seems to be about accountability alone. Probably the version about expressing attitudes is about accountability, but the version

about having attitudes is not about attributive responsibility. If this is correct, then the Standard View implies that a person, S, is negatively accountable for an action, A, if and only if, and in virtue of the fact that, it was morally wrong for S to perform A, and S had no excuse. It suggests a conceptual connection between moral wrongness and accountability.

If we can discover the conceptual connection between having core negative reactive attitudes and moral wrongness, then we can derive a conceptual connection between attributive responsibility and moral wrongness. Consider a view close to the Standard View, namely,

Standard View*: It is appropriate for reacting people to feel the blame-sentiment towards an agent, S, regarding S's performing an action, A, if and only if, and in virtue of the fact that, (1) S endorsed performing A, and (2) it was morally wrong for S to perform A.

The view eliminates the proviso about excuse and adds a condition about endorsing actions. Roughly, I take 'endorsing an action' to refer to the approval of an action and an willingness to perform the action. The reason is that attributive responsibility is about whether the action reflects the agent's mental state, regardless of whether the agent lacks information, is morally incompetent, or has no rational control over her action.

Accordingly, a person, S, is negatively attributively responsible for an action, A, if and only if, and in virtue of the fact that, S endorsed performing A and it was morally wrong

for S to perform A. It suggests a conceptual connection between moral wrongness and attributive responsibility.

I suspect that the Standard View* is correct. First, it is compatible with the idea that the natures of attributive responsibility and accountability are different. Second, it goes along with an influential view that attributive responsibility and moral wrongness are connected. This view is controversial (Scanlon, 1998, 2008), but it quite is plausible. The Standard View* may be vague, but I believe that it is at least on the right track.

In the next section, I will talk about some possible future work.

4. Future Work

Throughout this dissertation, I regard reactive attitudes to be cognitive emotions (emotions that presuppose thoughts), and I focus on negative ones. However, I have not studied any particular negative reactive attitude. In particular, I deem resentment, indignation, and guilt to be different sorts of blame-sentiments, but I have not studied what distinctive role they each play regarding morality. In future work, I may work on resentment, indignation, guilt, and shame individually. Furthermore, since resentment and indignation are other-regarding reactive attitudes, whereas guilt and shame are self-regarding ones, I may also work on the comparison between other-regarding and self-regarding ones.

Since the reactive attitudes I focus on are negative ones, I can definitely work on positive ones as a continuing project. Some might think that the story of positive reactive attitudes is just the flip side of the story of negative ones. However, things are not so

simple. Let's think of praise for simplicity.¹¹⁵ If an agent acts rightly at a minimum level (e.g., merely not violating any moral obligation), it seems inappropriate to praise her. But if she acts rightly at a higher level (e.g., fulfills a nontrivial moral obligation at a nontrivial cost), then her action is morally more desirable, and it seems appropriate to praise her. From these simple reflection, we can see that the appropriateness of praise is not simply a matter of acting rightly. It will be interesting to work on positive reactive attitudes and figure out how they connect with moral permissibility.

I suspect that the right account is like this:

Praiseworthiness: It is appropriate for reacting people to feel the praise-sentiment towards an agent, S, regarding S's performing an action, A, if and only if, and in virtue of the fact that, it was morally permissible for S to perform A, S performs A at a nontrivial cost to herself, and A is especially morally valuable.

With the accounts of both negative and positive reactive attitudes, we have more resources to account for supererogation and suberogation than with an account of negative attitudes alone.

Finally, the Standard View is merely about the conceptual connection between negative reactive attitudes and moral wrongness, but silent about the metaphysical connection between them. In the future, I may also work on the metaphysical connection.

I am inclined to think that the property of fitting negative reactive attitudes are grounded

¹¹⁵ Smith (1759) famously suggests that blame is the negative correlate of gratitude instead of praise. Here, I use praise as the placeholder of the positive correlate of blame. It's possible that the right account is about gratitude instead of praise.

in the property of moral wrongness, among other things. If this is right, and it is objective (independent of any one's mental state) whether an attitude is fitting, then an argument for moral realism (ontological sense) can be made and defended. Specifically, if the property of fitting attitudes is objective, and it is grounded in the property of moral wrongness, then the property of moral wrongness is also objective. Hence, we arrive at the claim that there are objective moral truths.

5. Conclusion

I have shown that the Standard View can help us understand the role fitting negative reactive attitudes play regarding notions such as moral permissibility, moral obligation, etc. I have also shown that the Standard View can help with the moral responsibility debate. Future work is needed on specific negative reactive attitudes, positive reactive attitudes in general, and the metaphysical connection between reactive attitudes and moral wrongness.

BIBLIOGRAPHY

- Boyd, R. (1988), "How to be a Moral Realist," in Sayre-McCord, (ed.), *Essays on Moral Realism*, Ithaca: Cornell University Press: 187-228.
- Brink, D. (1989), *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press.
- , (2008), "Mill's Moral and Political Philosophy," in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/mill-moral-political/>.
- Broad, C. D. (1928), "Analysis of Some Ethical Concepts," in *Broad's Critical Essays in Moral Philosophy*, David R. Cheney (ed.), London: George Allen & Unwin Ltd, 1971, pp. 63-81.
- Broome, J. (2008), "Comments on Allan Gibbard's Tanner Lectures," in *Reconciling Our Aims: In Search of Bases for Ethics*, Barry Stroud (ed.), Oxford: Oxford University Press, pp. 102-119.
- Butler, J. (1774), *Fifteen Sermons Preached at the Rolls Chapel*, London, M.DCCC. LXXIV.
- Coleman, J. and Sarch, A. (2012), "Blameworthiness and Time," in *Legal Theory*: 18: pp. 101-137.
- Dancy, J. (2004), *Ethics without Principles*, Oxford: Oxford University Press.
- , (2009), "Moral Particularism," in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/moral-particularism/>.

- Darwall, S. (2006), *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge: Harvard University Press.
- , (2010), "But It Would Be Wrong," in *Social Philosophy and Policy*: 27: 2: pp. 135-157.
- Enoch, D. (2011), *Taking Morality Seriously*, Oxford: Oxford University Press.
- Finlay, S. (2007), "Four Faces of Moral Realism," in *Philosophy Compass*: 2: pp. 1-28.
- Funkhouser, E. (2006), "The Determinable-Determinate Relation," in *No  s*: 40: 3: pp. 548-569.
- Gibbard, A. (1990), *Wise Choices, Apt Feelings*, Cambridge: Harvard University Press.
- , (1992a), "Pr  cis of Wise Choices, Apt Feelings," in *Philosophy and Phenomenological Research*: 52: 4: pp. 943-945.
- , (1992b), "Moral Concepts: Substance and Sentiment," in *Philosophical Perspectives*: 6: Ethics: pp. 199-221.
- , (1998), "Preference and Preferability," in *Preferences: Perspectives in Analytical Philosophy*, Cristoph Fehige and Ulla Wessels (eds.), Berlin: Walter de Gruyter, pp. 235-259.
- , (2003), *Thinking How to Live*, Cambridge: Harvard University Press.
- , (2005), "Truth and the Correct Belief," in *Philosophical Issues*: 15: pp. 338-350.
- , (2006a), "Pr  cis of *Thinking How to Live*," in *Philosophy and Phenomenological Research*: 72: 3: pp. 729-744.
- , (2006b), "Reply to Critics," in *Philosophy and Phenomenological Research*: 72: 3: pp. 687-698.

- , (2006c), "Moral Feelings and Moral Concepts," in *Oxford Studies in Metaethics*, Russ Shafer-Landau (ed.), Vol. 1: pp. 195-215.
- , (2008a), "Insight, Consistency, and Plans for Living," in *Reconciling Our Aims: In Search of Bases for Ethics*, Barry Stroud (ed.), Oxford: Oxford University Press, pp. 11-33.
- , (2008b), "Reply to Commentators," in *Reconciling Our Aims: In Search of Bases for Ethics*, Barry Stroud (ed.), Oxford: Oxford University Press, pp. 147-188.
- , (forthcoming), "Improving Sensibilities."
- Hieronymi, P. (2005), "The Wrong Kind of Reason," in *The Journal of Philosophy*: 102: 9: pp. 437-457.
- Hume, D. (1739-40), *A Treatise of Human Nature*, David Fate Norton and Mary J. Norton (eds.), Oxford: Oxford University Press, 2007.
- , (1751), *An Enquiry Concerning Human Understanding*, Stephen Buckle (ed.), Cambridge: Cambridge University Press, 2007.
- Hurka, T. (2010), "Moore's Moral Philosophy," in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2010/entries/moore-moral/>.
- Jacobson, D. (2011), "Fitting Attitude Theories of Value," in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/fitting-attitude-theories/>.
- Kant, I. (1785), *Groundwork of the Metaphysics of Morals*, Mary Gregor (trans.), Cambridge: Cambridge University Press, 1997.

- , (1788), *Critique of Practical Reason*, Mary Gregor (trans.), Cambridge: Cambridge University Press, 1997.
- Mackie, J. L. (1977), *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin.
- McDowell, J. (1985), "Value and Secondary Qualities," reprinted in *Mind, Value, and Reality*, 1998, Cambridge: Harvard University Press, pp. 131-150.
- Mill, J. S. (1861), *Utilitarianism*, George Sher (ed.), Indianapolis: Hackett, 2001.
- Miller, A. (2003), *An Introduction to Contemporary Metaethics*, Cambridge: Blackwell.
- Moore, G. E. (1903), *Principia Ethica*, Amherst: Prometheus Books, 1988.
- , (1912), *Ethics*, London: Oxford University Press, 1963.
- Nichols, S. (2004), *Sentimental Rules*, Oxford: Oxford University Press.
- Olson, J. (2004), "Buck-Passing and The Wrong Kind of Reasons," in *The Philosophical Quarterly*: 54: 215: pp. 295-300.
- Parfit, D. (2001), "Rationality and Reasons," in *Exploring Practical Philosophy: From Action to Values*, Dan Egnosson, Jonas Josefsson, Björn Peterson, and Toni Rønnow-Rasmussen (eds.), Aldershot: Ashgate: pp. 17-39.
- Prichard, H. A. (1937), "Moral Obligation," in *Moral Writings: H.A. Prichard*, Jim McAdam (ed.), Oxford: Oxford University Press, 2002.
- Rabinowicz, W. and Rønnow-Rasmussen, T. (2004), "The Strike of the Demon: On Fitting Pro-attitudes and Value," in *Ethics*: 114: pp. 391-423.
- Railton, P. (1983), "Moral Realism," in *Facts, Values, and Norms*, Cambridge: Cambridge University Press, 2003, pp. 3-42.
- Rawls, J. (1971), *A Theory of Justice*, Cambridge: Harvard University Press.

- Ross, W. D. (1930), *The Right and the Good*, Philip Stratton-Lake (ed.), New York: Oxford University Press, 2002.
- , (1939), *Foundations of Ethics*, Oxford: Oxford University Press.
- Scanlon, T. M. (1998), *What We Owe to Each Other*, Cambridge: Harvard University Press.
- , (2008), *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge: Harvard University Press.
- Schaffer, J. (2009), "On What Grounds What," in David Chalmers, David Manley, and Ryan Wasserman (eds), *Metametaphysics*, Oxford: Oxford University Press, pp. 347-383.
- Schroeder, M. (2012a), "Value Theory," in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2012/entries/value-theory/>.
- , (2012b), "The Ubiquity of State-Given Reasons," in *Ethics*: 122: pp. 457-488.
- , (2013), "State-Given Reasons: Prevalent, If Not Ubiquitous," in *Ethics*: 124: pp. 128-140.
- Shafer-Landau, R. (2003), *Moral Realism: A Defense*, Oxford: Oxford University Press.
- Skorupski, J. (2007), "Buck-Passing about Goodness," in *Hommage à Wlodek*: *Philosophical Papers Dedicated to Wlodek Rabinowicz*, T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonsson (eds.), Lund: Lund University.
- , (2010a), "Moral Obligation, Blame, and Self-Governance," in *Social Philosophy and Policy Foundation*, pp. 158-180.
- , (2010b), *The Domain of Reasons*, Oxford: Oxford University Press.

- , (2012a), "Pr écis of The Domain of Reasons," in *Philosophy and Phenomenological Research*: 85: 1: pp. 174-184.
- , (2012b), "Reply to Cassam, Olson, and Railton," in *Philosophy and Phenomenological Research*: 85: 1: pp. 210-221.
- Smith, A. (1759), *The Theory of Moral Sentiments*, London: A. Millar.
- Sturgeon, N. (1985), "Moral Explanations," in David Copp and David Zimmerman (eds.), *Morality, Reason, and Truth: New Essays on the Foundations of Ethics*, Totowa: Rowman and Allanheld, pp. 49-78.
- Strawson, P. F. (1962), "Freedom and Resentment," in *Proceedings of the British Academy*: pp. 187-211.
- Wallace, R. J. (1994), *Responsibility and The Moral Sentiments*, Cambridge: Harvard University Press.
- Wiggins, D. (1991), "A Sensible Subjectivism?" in *Needs, Values Truth*, second edition, Oxford: Blackwell, pp. 185-211.
- Zimmerman, M. (2008), *Living with Uncertainty: The Moral Significance of Ignorance*, Cambridge: Cambridge University Press.

VITA

Wenwen Fan was born in 1984 in Changsha, Hunan, China. She grew up with her family in Guangzhou. She attended Peking University, where she received her bachelor degree in philosophy. She got her master degree in philosophy at the University of Toronto and doctoral degree at the University of Missouri. She loves music, books, movies, and yoga.