

**ANSWERABLE AGENCY:  
A DEFENCE OF THE RATIONAL RELATIONS VIEW OF  
RESPONSIBILITY**

by

Bryan David Elliott

A thesis submitted to the Department of Philosophy

In conformity with the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

(April, 2012)

Copyright © Bryan David Elliott, 2012

## **Abstract**

The concept of responsibility is central to our moral practice. It is essential that any credible conception of responsibility appropriately tracks agency, that is, what an individual brings about as a consequence of their true self. The rational relations view is a contemporary theory of responsibility informed by recent work in the areas of practical reasoning and moral psychology, and offers a systematic and coherent account of the conditions for moral responsibility. However, the rational relations view faces challenges in deviating from concepts inherent in traditional responsibility theories, specifically the requirements of volitional control and a capacity for moral reasoning. I defend the rational relations view against these challenges, arguing first that volitional control is not a necessary (though sufficient) condition for ascribing responsibility, since such judgements track the evaluative commitment constitutive of our true selves, of which an agent need not be consciously aware. Consequently, against traditional volitional views of responsibility that characterize agents as authoring themselves, the rational relations view instead presents the practice as one of both self-discovery and self-creation. Next, I contend that no further capacity for moral reasoning beyond basic rational competence is required to be morally responsible, for being so simply consists in being answerable for one's evaluative judgements. I argue those who require a further moral reasoning capacity commit the traditional mistake of conflating moral responsibility with blame, which I take to be distinct. In exploring this distinction, I make use of psychopathy as exemplifying a class of agents who seem to lack a capacity for moral reasoning. The rational relations view is ultimately advanced as an account tracking an agent's true self without implying blameworthiness or praiseworthiness in its judgements of moral responsibility.

## **Acknowledgements**

I owe many people thanks for either directly contributing to this work or indirectly supporting it.

Foremost, I owe a great debt to my supervisor, Rahul Kumar. It has been my great privilege to work under his tutelage in both my undergraduate and graduate careers at Queen's, and this experience has been truly transformative. Without his guidance, this work would not have been possible.

My foray into philosophy came at the cost of time that would otherwise have been spent with friends and family, and so I must thank them for their forbearance: most especially, my wife Allison, and my parents, Jack and Karen. Without their support, I could not have been the kind of person who would take up this work.

## Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Chapter 1 - The Role of Conscious Volition.....	1
Chapter 2 - Necessary Capacities for Moral Responsibility .....	34
Coda.....	64
Bibliography .....	68

## Chapter 1

### The Role of Conscious Volition

Responsibility is an essential feature of our moral lives. Many take moral responsibility as setting the conditions which warrant expressions of praise and blame. When someone fails to meet a moral requirement, our criticism presupposes the responsibility of that individual for the failure; similarly, when we praise someone for a morally laudable act, we do so under the assumption that the act can be properly attributed to that individual. It is of critical importance to our moral practice that the agent “owns” the action, attitude, belief, or emotion being evaluated; indeed, the gravity of expressing blame or praise against an individual not culpable for it strikes us as deeply wrong and an affront to our own agency. Hence, a detailed and comprehensive theory of moral responsibility is needed.

Any theory of moral responsibility presupposes some corresponding criteria of moral agency. For, if we are to ascribe responsibility to agents for what they author, then we must have some account of how that authority is affected. This concern finds expression when we ask who can be considered a morally responsible agent. Often, we deem that young children, the severely mentally disabled or disordered, and the non-conscious<sup>1</sup> are not morally responsible agents. We have this intuition because such individuals seem to have a diminished degree of control over themselves, or since some such individuals, especially children, do not possess the experience or knowledge to understand the consequences of their actions. An account of responsible agency

---

<sup>1</sup> I have in mind here sleepwalkers or perhaps those in fugue states, though perhaps the latter is better captured by the mentally disordered category; however, individuals in such a state generally retain their normal rational and physical capacities though they lack any conscious awareness, or later memory, of their activities.

will specify the conditions an agent must meet in order for ascriptions of responsibility to be imputed against her. A common criterion is that the agent must have control over the object of assessment in order to be responsible for it. This Control Principle, the idea that an agent “cannot be morally assessed for what is not their fault, or for what is due to factors beyond their control” (Nagel, 25), will have varying targets depending on the concept of agency being deployed, but reflects the generally accepted premise that our agency is expressed in our authority over the certain aspects of ourselves subject to such control.

The concept of agency underlying a theory of moral responsibility will have two further consequences for our moral practices. First, it will identify what features of an agent’s life can be the focus of moral assessment. For example, in keeping with the aforementioned Control Principle, a view of agency that sees us as active with regard to our actions but passive with regard to our emotions (i.e., such feelings simply assail us without relation to our agency) would as a result hold us morally responsible for our actions but not for our feelings: we would take Jones to be responsible for striking his wife, but not responsible for being angry with her. Second, the conception of agency will identify excusing conditions where moral responsibility cannot be imputed or must at least be attenuated. Perhaps Jones had been hypnotized or otherwise conditioned to hit his wife when he was angry with her; perhaps Jones was drunk and acting out of character; or, perhaps Jones was morally ignorant of the fact that such domestic assaults are wrong. The viability of each preceding excuse will turn on the authority that a particular conception of agency takes us to have in the given circumstances. Put simply, the conception of agency deployed will answer the questions of “what” we can be responsible for and “when” we can be responsible for them. Obviously then, substantial and centrally important features of any

theory of moral responsibility will depend upon the particular conception of agency upon which it is built.

Despite the centrality of responsibility to our moral practices, there persists a debate regarding the conditions that must obtain for ascriptions of responsibility to be made. A traditionally dominant view is that some voluntary action on the part of the agent is a necessary condition for imputing responsibility. Such volitional views point to our past choices, identification with present attitudes, or future alterability as the agentic capacities that ground responsibility judgments. Generally, those who take a view of ethics as *action-guiding* are attracted to such volitional views, for these track the choices, the actions and omissions, with which they take morality to be centrally concerned. Thus, an assumption about the morally relevant characteristics of our agency, here the ability to choose, is taken to inform the appropriate conditions for the ascription of responsibility.

An alternative view of responsibility has been recently advanced, initially by Tim Scanlon and more fully by Angela Smith. This view holds that we are responsible not simply for explicit choices, but also for our attitudes, beliefs, and emotions insofar as they are rationally connected to our evaluative judgments. This rational relations view of responsibility casts a wider net than the volitional view. However, like the volitional accounts, the rational relations view is also founded on a particular conception of how our agency operates. Consequently, as the conception of agency endorsed by Scanlon and Smith extends beyond our explicit choices, an agent is subject to bearing responsibility for a broader array of judgements than just their conscious decisions. This is appealing, as many of our intuitive moral judgements presuppose responsibility where there has been no explicit choice on the part of the agent, and the rational relations view can account for and justify such intuitions.

In this chapter, I shall examine in detail both the volitional view and rational relations view. My aim shall be to illustrate the varieties of volitional view on offer, and how each ultimately fails to capture important aspects of our moral practice. The rational relations view is advanced, and its ability is demonstrated not only to address the shortcomings of the volitional views, but to account more robustly for the ascriptions of responsibility reached by such choice-focused accounts. The nature of our agency is a continuing theme underscoring the discussion of all the accounts under consideration.

### **Volitional Views of Responsibility**

The volitional view holds that some explicit conscious choice by the agent is a necessary condition for responsibility. This reflects the widely held view that perhaps the most essential feature of responsible agency is the control we are able to assert over what we do. In contrast to other animals, human beings are able to think about what course of action they will take, and further are able to reason, based on their experience, about what the likely consequences of these actions will be. This ability to reflect upon reasons and to predict likely outcomes is said to confer upon us responsibility for what follows. We are responsible where our control is not constrained. Such constraint might manifest itself through exotic means such as hypnosis, brain stimulation by rouge neuroscientists or Cartesian demons, or through more mundane means such as coercion or duress. Certainly, in any of the aforementioned cases the claim by the agent that “I had no choice!” strikes us as a plausible excuse. Likewise, our responsibility for the consequences that follow from our actions seems attenuated when these outcomes could not reasonably have been anticipated by the agent. When I ring a doorbell, unaware that inside Jones is adjusting the wiring, I am not held responsible for the resulting shock that Jones receives, for I could not have reasonably known that my action would lead to such an outcome. The concept of reasonability is

central here, for some agents go about the world in greater ignorance than others; thus, what one ought to reasonably have known sets a standard that responsible agents ought to achieve, but often do not. This normative epistemic standard fits well within the volitional view of responsibility, for when an agent fails to meet it that failure can often be traced to a choice. Thus, even actions where an agent did not know the likely consequences, but should have, can be reduced to an explicit choice. These explicit choices may manifest themselves in several ways, specifically prior choice accounts, present identification accounts, and future alterability accounts.

Prior choice accounts take the view that we are responsible for the consequences that issue from our past decisions. Traditionally, this has been the dominant account of moral responsibility since Aristotle<sup>2</sup>. Such prior choice views depend on being able to give an account of how the agent's past decisions gave rise to the outcome under assessment. The outcomes which can be subject to such tracing are diverse. Someone who in the past made a conscious decision to begin reading racist literature, and subsequently developed strongly racist attitudes, would be responsible for these resultant attitudes under this view. A prior decision to decline treatment for an injury that subsequently becomes infected would render the agent responsible for the consequence of his action. Choosing to ignore particular pieces of evidence that contribute to one holding false beliefs similarly renders one responsible. If an attitude, belief, or state of affairs can be traced to the prior choice of an agent, then, excepting the excusing circumstances common to all volitional views, the agent bears responsibility for the respective attitude, belief or state of affairs. Again, this responsibility is grounded in the agent's explicit decision, reflecting, *ex hypothesi*, her deliberative control and outcome expectations in choosing what to do. We are

---

<sup>2</sup> See his *Nicomachean Ethics* III.1–5.

responsible just in virtue of potentially having done otherwise and by knowing what would reasonably come about. Responsibility, on the prior choice account, is thus “after the fact”. An agent makes a choice, and we must wait to see what the ultimate morally relevant results will be.

The general acceptance of this account has led to its expression in other areas of philosophy, most notably political philosophy. Individuals who fare poorly in society are often deemed to be responsible for their position in virtue of their prior choices. Perhaps they passed up opportunities for further education, or elected not to work very hard, or made foolish investments with their money. In such cases it is assumed the individual could have chosen to do other than she did, and that she ought to have had a sufficiently good understanding of the consequences of the choices with which she was presented. The prominence of the prior choice view can be seen in the immediacy with which politicians appeal to it as recourse against particular social criticisms.

In addition to the backward-looking prior choice account, there is a view that grounds moral assessment in the present identification of an agent with her fundamental desires. This view is implicit in Harry Frankfurt’s account of agency. Frankfurt takes us to be primarily moved by first-order desires, these being motivationally efficacious states. Such a first-order desire might be to eat or drink, or to be loyal or brave. Beyond these first-order desires are second-order desires, which consist in wanting to be possessed of certain first-order desires. Such second-order desires could include desiring that one’s appetite return (i.e., that one have the desire to eat), or to desire that one act in a loyal manner or courageously. When one acts, one does so on a first-order volition, an act of will that aims at fulfilling some first-order desire. But, while these first-order volitions are expressive of an agent’s will as an actor<sup>3</sup>, they are not expressive of the agent’s will

---

<sup>3</sup> Frankfurt uses the term “wanton”, who, while possibly rational, simply acts without reflection. I use the term actor to avoid the negative connotations of “wanton” without necessarily imparting the moral dignity I take persons to have.

as a person. Frankfurt bases this distinction between actors and persons upon second-order volitions, these being the desire that some particular first-order desire (or set thereof) become the agent's will. Acting on a first-order volition to realize a first-order desire for safety, an agent might flee a dangerous scene; alternatively, the agent may also possess a first-order desire to be brave, and a second-order desire to want to be brave, and thus form a second-order volition that one act from the desire to be brave and consequently remain at the perilous venue. For Frankfurt, an agent acts freely when her second-order volitions are what determine her actual will. Moreover, since forming a second-order volition essentially involves endorsing particular first-order desires as those which ought to determine one's action, responsibility thus tracks how closely our actions were predicated on attitudes we identified with when we acted.

Under the present identification variety of the volitional view, the fact that the agent endorses a racist attitude is sufficient for it to be attributed to her for the purposes of moral appraisal; likewise, her disavowal similarly dispels moral responsibility. This deferral of responsibility through disavowal may seem evasive, but in fact reflects a fundamental tenet of Frankfurt's conception of agency. Namely, on Frankfurt's view our first-order desires are states which simply assail us unbidden. It is only through identifying with such desires through second-order volitions that we can take ownership of the attitudes and actions which ensue. To hold someone responsible for their first-order desires or results of their first-order volitions consequently violates the Control principle, since such states and wills are not chosen by the agent.

The present identification view imputes responsibility for states and acts we may not have brought about but which we nevertheless choose to endorse. But not all unbidden states or actions are ones which we will opt to identify with, or even be able to influence through our agency. Consequently, some posit that one is responsible only for attitudes that are susceptible to revision

by the agent. On this future alterability view, what is essential for an attribution of responsibility is that the attitude be subject to control through the agent's future choices. Like the present identification view, the future alterability view is indifferent to how the attitude came about, but it differs from the present identification view in that endorsement of the attitude is not a necessary condition for an ascription of responsibility; rather, it is the susceptibility of the attitude to volitional revision that determines what attitudes we can be held responsible for. Thus, while one might discover oneself possessed by a racist attitude towards some group, presumably such an agent could work to suppress that attitude, to "will it away" through their control over their own psychology. One would only escape being responsible for holding this objectionable disinclination if the attitude were impervious to alteration, perhaps by reason of some mental disorder. Here again, the lack of control by the agent over the feature being morally assessed is the excusing condition for ascription of responsibility. Future alterability views thus assume the agent possesses a high degree of control over their psychology, else it would lose credibility in the face of excuses that "*I couldn't help myself*".

Any volitionalist account will depend on some connection to past, present, or future voluntary choices to make attribution of an action, attitude, or belief to an agent appropriate for moral appraisal. As evinced by each of the three views just proffered, the conception of agency at play in a particular view defines the nature of the choice which grounds responsibility. The epistemic situation of the agent may be relevant, as in the prior choice view. The more reflective endorsement of the present identification view contrasts with the more mechanical practical reasoning of the prior choice view, and seeks to promote the idea of self as author. Even the future alterability view supposes that an agent critically self-evaluates herself and consequently chooses to improve or not. Common to all these conceptions of agency is the idea that choice

reflects judgements about value. On the prior choice view, an agent might choose for good reasons to act in a manner that aims at a good outcome, but through some unforeseeable set of circumstances instead ends up causing harm. Such an agent is not judged morally responsible, just because the nature of his action aimed at the good, and there was nothing in his self-governance that was faulty. On the present identification view, the agent's formation of a second-order volition represents her evaluation of particular first-order desires as good, and we hold the agent responsible for this endorsement and the value judgement it entails. Value judgements are most pronounced in the future alterability view, wherein the agent is exactly deciding what sort of attitudes and actions it is good for her to embody.

Insofar as ethics seeks to answer the question, "How should I live my life?", and inasmuch as this turns largely on judgements concerning value, volitional views of responsibility hold tremendous appeal. Indeed, it is no wonder that such accounts have held sway over moral theory for so long, when we consider the importance we place on self-ownership and the role played by cognizant choice in moral discourse. The former provides an intuitive theoretical grounding whereas the latter gives us practical examples that reinforce the applicability of the view.

The idea of self-ownership or self-mastery has a long philosophical history. Aristotle stressed the importance of self-mastery throughout his moral theory<sup>4</sup>. More recently, we find this concept taking a central role in Kant's moral theory, wherein he takes the autonomy of the will to be an essential condition for moral action<sup>5</sup>. Intuitively, we desire control over our lives, as constituted by our actions, attitudes, beliefs, and emotions. We take it as essential that we vote for our chosen candidate and not another; we dislike a racist on the basis of our negative assessment of his view; we believe in atoms since we accept the science which posits them; and, we feel joy because we

---

<sup>4</sup> See especially his *Nicomachean Ethics* VII.

<sup>5</sup> See his *Groundwork for the Metaphysics of Morals*, Chapter 2.

judge the situation to warrant it. Conversely, it strikes as contrary to our agency when such aspects of our lives oppose our conscious desires. It is an affront to our conscious selves when a decision to avoid sugary desserts is overturned by akratic action; when we seek the approval of those we judge unworthy of our friendship; when we cannot shake beliefs that are undermined by compelling evidence; and, when we experience sadness or anger absent cause. To be sure, our consciousness ensures that we are inherently self-reflective. We are so immediately aware of our actions, attitudes, beliefs and emotions in virtue of our consciousness that it is difficult to divorce ourselves from them; they are not merely events that happen to us, rather, they are constitutive of who we are. Given the usual conscious control over these aspects of our mental lives, and the transparency of our explicit agency to these aspects as the constitutive components of our lives, it is very appealing to identify conscious choice as the locus of responsibility. As conscious agents we clearly own our choices, and the resultant actions, attitudes, beliefs, and emotions just are our consciously endorsed choices. Thus, we are directly responsible for such aspects of ourselves in virtue of these conscious choices.

The practical import of volitional views is evident in our moral discourse. By having responsibility track explicit conscious choices, moral discourse is facilitated through the unproblematic identification of objects for moral assessment. We do not need to guess after the evaluative judgements held by individuals, since we are concerned with their conscious, and thus implicitly knowable, choices<sup>6</sup>. Moreover, the force of our moral criticism seems strongest when pressed against an agent's conscious choices. Expressing moral criticism against someone when they did not consciously make a choice strikes us as *prima facie* wrong. We need only consider the well-used example of someone who is pushed into another; clearly, they are not responsible

---

<sup>6</sup> Obvious, one might choose not to cooperate in such moral assessment; however, I only wish to illustrate here that volitional views have the advantage of concrete objects of moral assessment.

for any injury that occurs, and this is in virtue of their lack of choice in the action. Such examples must be qualified though, lest license for omissions be granted. Omissions, then, must be viewed as choices not to act. Negligence is a prime example here: one *chooses* not to take a required action, and this choice grounds the ascription of responsibility for an action *not* taken.

The volitional view of responsibility is appealing in that it ascribes responsibility as a result of our conscious activity. We are responsible as a result of antecedent choices, or our present judgements, or for that which we can affect. Yet, we do not always respond to what people explicitly chose to do or believe. As many philosophers have pointed out, we also react to other's spontaneous, unreflective, and implicit attitudes. Moral assessment does not track simply the conscious decisions made by an agent. Instead, moral assessment is concerned with the "whole person": what they notice and what they neglect, what they see as reasons, and their spontaneous reactions, all in addition to their consciously lived choices and decisions. Consequently, the volitional view misses a vast segment of our moral life. Let us consider some of the ways how.

### **Weaknesses of the Volitional Views**

What we notice and fail to notice in our environment is not, generally speaking, something which is under our volitional control. While we may intend to be more or less attentive, what we actually perceive is not something determined by will. Nevertheless, we place great significance on what a person does or does not notice, for this certainly seems to express an affinity for what they value. We arguably take note of things that we place great value in, whereas we often neglect those things which leave us cold. The role this plays in our relationships is illustrative. If I claim to value someone's friendship, yet consistently fail to take note of things that would benefit this person or revise behaviours that I understand annoy her, then the actual value of this friendship for me is thrown into question. The normative standards that define friendship demand

that such considerations about my friend figure in my evaluative judgements; failing to notice implies that I do not take the normative standards to apply, i.e., that I do not actually consider this person my friend. Similarly, consider someone who claims to be unbiased, yet, on meeting a new person, always first notes their race. While this individual may not act in any overtly racist manner, the fact remains that a person's race is, for him, a predominant feature. Thus, what he explicitly claims to be unimportant to him does in fact have evaluative significance for him. These examples demonstrate that what we notice and fail to notice can have a moral significance for our relationships with others, despite their insusceptibility to immediate volitional control.

The nature of this significance might be questioned, as different agents with similar evaluative commitments will certainly notice different things pertaining to that commitment in a given circumstance; moreover, when we are tired or under stress, we may fail to take notice of considerations that bear on even our most deeply held values. But the principle here being advanced is neither as particular nor demanding as these objections would require. Rather, it makes a more modest claim about our moral psychology, that "if one judges some thing or person to be important or significant in some way, this should (rationally) have an influence on one's tendency to notice factors which pertain to the existence, welfare, or flourishing of that thing or person" (Smith, 2005). Thus, we do not expect that different agents with similar evaluative judgements will perceive the *exact* same things, but will instead exhibit an affinity for perceiving various considerations that bear on their object of importance. And, since the principle only claims an influence, it is understandable how one's failure to take notice under certain exceptional mental circumstances would not reflect on one's evaluative commitments.

While failing to notice has thus far been used to illustrate morally blameworthy behaviour, it is worth taking account of how such failings can be praiseworthy. Someone who regularly fails to

take note of minor trivial faults in others should be applauded, for such behaviour reflects a genuine judgement that such faults are not of any moral importance. This point can be drawn out to contrast with a contention made by Kant. Whereas Kant claimed that the benefactor moved to generosity by his giving nature does not perform a morally worthy act when giving alms, on the view being advanced the benefactor's generous nature would be morally creditworthy. The distinction here is one between what an agent is like (her evaluative commitments) and what an agent sees as reasons for an action (what occurs to her). Generosity is something that can be attributed to the benefactor: he preferentially sees opportunities where his giving would benefit others. Presumably, the considerations that led to the perception of an opportunity to give will inform the consequent action as well. But this is not necessarily so, as the giving might be done to maximize benefit, to impress a peer, or to shame the recipient; each of this additional considerations moves the act of giving further along a spectrum of moral repugnance. The generous benefactor can thus be said to possess the morally praiseworthy virtue of generosity, while remaining answerable for the considerations that moved him to give.

What considerations strike us as reasons is another process not normally under our voluntary control. Yet, the fact that we do see particular things as reasons does have moral significance. As in the preceding example, seeing that an act of generosity would shame the recipient and considering this as something that would count in favour of the action, even if the consideration is ultimately rejected, reflects something negative about the benefactor. Such mental occurrences differ from noticing in that the former concerns the internal operation of our reason whereas the latter deals more with the perception of considerations. Thus, forgetting can be viewed as a species of the former type. Forgetting an important appointment, barring the aforementioned mitigating circumstances, indicates that one did not actually take the appointment to be important.

In contrast to all the occasions when one's faculties bring to the fore considerations that bear on what one values, forgetting seems to imply that the subject does not really matter to the agent. This is what is disturbing about the friend who forgets a lunch date and the benefactor who considers giving to shame: both imply through their deliberative (in)action that the other does not matter to them. When we are called to answer for such occurrences or lapses, there are two reasons that motivate our response. Primarily, we seek to reassure others of our avowed evaluative commitments. We tell the stood-up friend that we do, in fact, value their company and enjoy being out with them. The benefactor expresses regret for considering an action that would harm another unjustly. Secondly, such acknowledgements serve to reinforce in ourselves the evaluative commitments we explicitly endorse. Such acknowledgements constitute a kind of self-criticism that seeks to cement the valuations we consider important in our lives<sup>7</sup>.

Our spontaneous emotions and other such attitudinal reactions are another class of actions which, while lacking immediate voluntary control, still seem to matter morally. People or situations often elicit a wide variety of emotions and attitudes, such as amazement, disdain, gratitude, jealousy, laughter, and regret. These involuntary responses often convey a great deal of expressive significance. Our amazement often signals that we did not believe an individual capable of some action or feat. A swelling of disdain when meeting a new person may reflect a hidden prejudice. Similarly, reacting to news of someone's misfortune with spontaneous laughter certainly conveys great significance as to how you view that individual. This example can be contrasted with regret, especially the variety discussed by Bernard Williams<sup>8</sup>. The lorry driver, who through no (voluntary) fault of his own runs over and kills a child, will be overcome by

---

<sup>7</sup> This point echoes the central idea behind the future alterability view. Answerability here is both a response of our considered evaluative judgments and a psychological assertion that these be internalized.

<sup>8</sup> Bernard Williams, "Moral Luck", in *Moral Luck* (Cambridge: CUP, 1981).

regret despite the reassurances of others that he has done nothing wrong; moreover, we would find it odd if the lorry driver were to easily shrug off the incident. Williams uses this observation to illustrate a certain coldness in the prior choice view, which would claim the lorry driver bears no responsibility for the child's death and ought not worry about it.

All these examples illustrate the manner in which our immediate reactions can express our truly held values. Such expressions have an obvious moral significance. To feel contempt for someone implies that you judge them not to be worthy of your respect. Regret indicates an acknowledgement that something of value has been lost. Consequently, many of our spontaneous and immediate reactions exhibit a conceptual connection to our having, or being inclined to have, certain judgements. Such judgements are contingent on our evaluative commitments, and thus are something that can be ascribed to us and for which, in certain circumstances, we can be asked to answer for. In many cases, these reactions may reveal something surprising or unpleasant. But most often, they simply serve to confirm our explicitly endorsed attitudes and commitments.

The examples of what we notice and neglect, what occurs to us in our deliberations, and our spontaneous reactions illustrate that more than our simply explicit choices and decisions have moral significance. The volitional view requires active engagement and dismisses our passive experiences as irrelevant to moral assessment. But as the examples above demonstrate, we discover moral significance in both the active and passive aspects of mental life. The significance concerning us here, the type that matters for ascriptions of responsibility, appears to attach to judgements held by an agent as expressed by their evaluative commitments. Restricting the criterion of responsibility to the explicit choices and decisions of an agent, as advocated by the volitional views, ignores a plethora of morally relevant facts about ourselves which matter greatly

to both self- and other- assessment. Consequently, such theories “leave us with an impoverished conception of moral personhood” (Smith, 2005, 263).

### **The Rational Relations View of Responsibility**

Responsibility thus seems to track, in contrast to traditional volitional views, not our explicit choices but rather the judgements we hold, insofar as they reflect our evaluative commitments. As noted previously, the choices that ground responsibility in volitional views are ultimately choices about value; however, as the foregoing discussion revealed, our evaluative judgements are evinced in more than just our explicit voluntary choices. More needs to be said on the nature of these judgements and their rational connection to our evaluative commitments. As well, there are several classes of judgement which might seem to conflict with what we consider appropriate subjects for moral assessment. In addressing these topics, Angela Smith’s rational relations view provides what I believe to be a superior account of responsibility. Put briefly, the rational relations view holds that moral praise or criticism concerns certain judgements held by a person that are reflected in their actions, attitudes, beliefs, and emotions, which are appropriately related to their evaluative commitments. Before examining the view itself, it is worthwhile first considering the conception of agency to which this view subscribes. As will become evident, there is a direct linkage between what the rational relations view deems one can be responsible for and the characterization of agency which denotes the boundaries of one’s affective powers over oneself.

Smith’s view is most strongly influenced by Scanlon’s rationalistic conception of agency and view of our practical reasoning. In giving an account of responsible agency, Scanlon asserts that such persons are conscious, rational, and embodied. They are conscious in that they are the subject of a stream of thoughts, perceptions, and experiences. Though this stream is segmented by

periods of sleep and other episodes of unconsciousness, it remains unified through the constancy of its content, the referential coherence of this content, and its theoretical causal basis. The first aspect of this unity is expressed in “the high degree of continuity in my cognitive and affective reactions: in what I like and dislike, in what I believe and reject” (Scanlon 2002, 168). The second aspect finds expression in “the way in which elements of my conscious life refer to each other, as when it seems to me that I am remembering a past experience of decision or carrying out a previously formed intention” (Scanlon 2002, 168). As pertains to our active agency, consciousness as described is essential. We would feel alienated from ourselves if our affinities and beliefs shifted dramatically in the intervals between going to sleep and awakening. And, while we may be fickle and frequently change our plans, this vacillating tendency must exhibit referential coherence else we lose ourselves in the changeable sea of inconsistent intentions. Yet, despite its importance for our active agency, consciousness is not a necessary condition for attributions to be made. Though an agent is unaware of particular attitudes, they may still be appropriately attributed to the agent should they be the best explanation of her behaviour, both in what she thinks and fails to think.

Responsible agents are rational in that they are able to evaluate reasons and consequently form judgement-sensitive attitudes<sup>9</sup>. These judgement-sensitive attitudes are not always conscious. Rather, they are attitudes that we adopt as a result of the reasons we evaluate as counting in favour of holding them. Insofar as we are rational, we will adopt such attitudes when we judge there to be sufficient reason to justify them, and we will cease to have them should we judge the relevant reasons to be insufficient. Our rational faculties present three features that are germane to

---

<sup>9</sup> Scanlon’s judgement-sensitive attitudes are equivalent to Smith’s evaluative judgements / evaluative commitments. All three terms can be used interchangeably, as they all reference an agent’s valuation of reasons.

the question of responsible agency. First, we experience perceptions of considerations that strike us as reasons. We notice that someone is burdened, which perhaps is a reason to assist him. It occurs to us that our friend's birthday is today, which gives us reason to call with best wishes. Second, we are able to assess considerations to determine whether they are in fact good reasons. Perhaps the burdened individual is so encumbered for physical training purposes and would not welcome our assistance, and perhaps our friend hates birthdays and would resent our phone call. Finally, we make up our minds as to what attitudes to adopt, whether these are intentions, beliefs, or emotions, when we judge there to be sufficient reasons counting in their favour. These *seemings*, *assessings*, and *optings* are normatively and descriptively related to our lives. They are descriptively related, in that if one did not have these rational features, one could not claim to be a rational agent. They are normatively related in that they ought to determine what we do, though we often fall short of the perfectly rational standard.

The rational features at play in our agency are expressed not only consciously, but also in what we do without thinking. What we evaluate as reasons is expressed in what we perceive or fail to perceive, what occurs to us or fails to occur to us, and in what we spontaneously do and believe. The conceptual linkages between these features guarantee their operation independent of our conscious awareness.

Based on this sketch of rational agency, one might envision two extremes of attributability. On one hand, we might posit that every feature of our mental lives, be they attitudes, explicit and implicit beliefs, intentions, emotions, and even somatic sensations, are attributable to an agent. This "gross literal truth" view would also encompass induced mental states (those brought about by rogue neuroscientists as well as Cartesian demons) in addition to mental states produced through normal means. Consequently, this view seems to cast too wide a net to account for our

considered judgements about responsibility. Clearly, one should not be held responsible for mental states brought about by brain stimulation or demonic alteration, any more than one should be called to account for having a headache. At the opposite extreme, an agent is only viewed as being responsible for her consciously made choices. Much of the appeal of this approach lies in the concrete choices that can be identified on which responsibility supervenes. This view also has the illusory appeal of tracking the weight of our judgements in cases where possibilities are constrained. We judge the person who assists a thief under duress as less culpable than one who helps willingly. But this view is ultimately unsatisfactory as well, for it takes too narrow a view of what can correctly be attributed to us. Cases of negligence provide an entire class of examples wherein no conscious choice has been made yet agents clearly bear responsibility for their oversights. Additionally, in general practice we tend to hold agents accountable for their spontaneous reactions. Individuals who unreflectively blurt racist comments are considered responsible for the outburst and viewed in consequence of it as holding particular attitudes.

This last point suggests the grounding of responsible agency in Scanlon's view. The attitudes we hold ought to be sensitive to our judgements in virtue of our rational faculties. When these attitudes do not correspond to our explicit judgments, then we ought to be able to modify them through conscious and reflective effort. The evaluative commitments expressed by our attitudes, whether explicit or implicit, are thus the appropriate subject for attributions of responsibility. Because we are embodied agents, we are prone to physical malfunctions that may impair our consciousness or our rational faculties. Consequently, we sometimes hold irrational attitudes that might not be appropriately attributed to us for the purpose of *moral* responsibility. Likewise, our consciousness may sometimes be subverted by mental illness such that we are not able to monitor and modify our judgements, and here as well attribution may be inappropriate for *moral* criticism.

Such qualifications may seem an evasion, but they simply reflect the material contingency of our agency in the physical world as embodied actors. Thus, ascriptions of responsibility track the immutable aspects of agency, our consciousness and rational faculties, while maintaining a sensitivity to the myriad issues raised by being in the world. To be a responsible agent, in Scanlon's view, is to be possessed of a normal, coherent psychology characterized by consciousness and rational features that enjoy strong linkages between *seemings*, *assessings*, and *optings*.

The role played by our rational faculties in Scanlon's account of agency opens up space for assessment of involuntarily<sup>10</sup> held attitudes. The judgements at play in the rational relations view are not only consciously held decisions and beliefs, but also include a broader set of tendencies to view certain aspects of the world as having evaluative significance. Such judgements, assessed holistically, express our take on what is valuable, what we consider important, and on what we place significance. As such, these judgements need not be consciously held by an agent. As Smith notes, we often only come to discover certain of these judgements when faced with a situation on which they bear. One might not realize how much they value an occupation that provides challenges over financial remuneration until faced with a dilemma that demands such a choice. Such practical necessities, situations where an agent must necessarily perform a particular action, are strongly reflective of who we are, despite the fact that we may not have known such important truths about ourselves prior to facing the situation that drew them out. The evaluative commitments that such judgements reflect are arguably more constitutive of who we are than our actively authored avowals. These judgements give moral criticism its special force, since they often have serious implications for our relations with others. Insofar as we are motivated to justify

---

<sup>10</sup> Here though involuntary simply means not consciously chosen, where the operation of our rational faculties was not affected by any foreign force.

ourselves to other agents, the judgements expressed by our evaluative commitments will be the currency of the exchange.

Saying that a mental state reflects an agent's evaluative commitments is to claim that if one sincerely holds particular evaluative judgements then such mental states should occur as a consequence of the normative link between the judgment and the mental state in question. Smith gives the fear of spiders as an example. If one truly judges that there is nothing about spiders that gives one reason to be afraid of them, then one ought not to have the emotional response of being fearful of them. A similar phenomenon is at play in belief, where if I truly judge the evidence to be in favour of a certain proposition, then I should subsequently form a belief that the proposition is true. Likewise, if I judge I have undefeated reasons to perform some action, then I should form the intention to execute that action. However, our rationality does not always meet this normative ideal, and often the connection between judgements and mental states will fail. Yet, such failures do not demonstrate a weakness in the view being proposed, but rather illustrate the fact that human beings, as embodied agents, are not always fully rational.

Such failures are sometimes cases of true irrationality, but more often, agents may simply be holding substantively incorrect judgements. Returning to the spider example, the continuing state of fear is more likely to reflect a (mistaken) belief that spiders are extremely dangerous or threatening. Thus, we can criticize the person in this case for having a faulty belief. Should the person come to sincerely accept that spiders are not dangerous or threatening (or objectionable in any other way), then we would expect the associated fear to desist. The continuance of such a state indicates serious malfunction of one's rational faculties. Such truly unreasoning cases of

arachnophobia do appear to be cases of actual irrationality, and as such escape ascriptions of responsibility<sup>11</sup>.

The failure of our mental states to sometimes properly reflect underlying evaluative judgments imparts an important qualification on the rational relations view. It presupposes that agents are possessed of a coherent psychology wherein the rational connections between her mental states and the evaluative commitments which she accepts obtain. Such a connection is necessary for ascriptions of responsibility, for the rational relations view takes only the evaluative commitments of an agent as the legitimate subject of moral appraisal. Thus, when mental states become unlinked from the evaluative judgments, they no longer are reflective of what the agent is committed to in a way that makes criticism or praise intelligible. Calling the truly irrational arachnophobe responsible for the state of fear seems wrong, insofar as her evaluative commitments do not support such a reaction. Asking her to answer for her fearfulness similarly runs up against the fact that she can see no reason to be fearful. Finally, ascriptions of blame seem out of place, as there is no way for the agent to modify her mental state; indeed, the very problem here is that the mental state is not responsive to her rational processes.

The importance of a coherent psychology to the rational relations view is underscored by considering the case of “implanted” mental states. We can imagine such implanted mental states as being brought about in an agent by a third-party, perhaps by hypnosis, brain stimulation, or pharmacological methods. For the agent, these implanted mental states cannot be differentiated from their native ones. If pressed, the agent would avow such exogenous states in the same manner she would for any of her endogenous judgements. What distinguishes these implanted attitudes is that they lack any connection to the agent’s evaluative commitments. Whereas

---

<sup>11</sup> Such an agent would still be open to criticism for being irrational, but this criticism would not be *moral* in nature.

normally her judgements come about as her rational faculties express her evaluative commitments, in implantation cases these rational faculties are bypassed altogether. This connection can sometimes be tenuous, and our resulting ascriptions of responsibility equally so. Consider the fleeting impressions that assail us daily. Such random thoughts or images may be violent or sexual in nature. When these occur infrequently, they seem insubstantial from a moral perspective, perhaps instead being indicative of the prevalence of such imagery in our present culture; but, when one is constantly beset by such impressions, there then seems to be a link to the agent's evaluative commitments. In this latter case, we may judge the agent responsible for the attitudes expressed by these thoughts<sup>12</sup>. Other aspects of our mental life, such as somatic sensations and appetitive desires, lack an appropriate connection to our evaluative commitments. Being in pain or experiencing thirst is expressive of one's physical rather than psychological state. Responsibility, on the rational relations view, depends upon the reliable relation between an agent's evaluative commitments and their ensuing mental states brought about by the normal functioning of their rational faculties as expected in a coherent psychology.

It is important to note that while the rational relations view ascribes responsibility for various mental states, it is not an aretaic view. It might support the claim that an attitude is properly attributable to an agent, in virtue of the rational linkages it bears to her evaluative commitments, yet it still assumes answerability, if only in principle. Any attribution of an attitude (e.g., A wants to  $\Phi$ ) implies, in virtue of the rational nature of these attitudes, that an answerability demand can also be made (e.g., A thinks p justifies  $\Phi$ ing). The principled answerability of the rational relations view is thus a necessary consequence of its theoretical underpinnings. Moreover, it

---

<sup>12</sup> More would need to be said about the agent's overall psychological health here. Many disorders can result in such unfortunate mental preoccupations. In such cases, the coherent psychology presupposed by the rational relations view does not obtain.

reflects a normative assumption of the view, that agents ought to be able to give an account of their actions, attitudes, beliefs, and emotions, in light of reasons they take to respectively justify them.

The rational relations view provides a superior account of responsibility to traditional views, particularly those of the volitional variety. It can account for all the cases identified by volitional views, for in such cases explicit choices and decisions are generally transparent to the agent's evaluative commitments. Barring mental disorders, our conscious deliberations are exemplary instances of our rational faculties producing judgements that reflect our evaluative commitments, and as such are exemplary instances of situations in which we are clearly responsible. But the rational relations view goes further, and can also account for the implicit and involuntary attitudes that have significance for both moral and non-moral assessment. Utilizing the view, we can speak on both the wrong extant in an individual's violent obsessions and the arachnophobe's unreasoning fear; the former has great import for expectations of the individual's behaviour towards others, and the latter demonstrates a failure of the individual's rational faculties. Moral assessment on the rational relations view concerns itself with the evaluative commitments of the agent rather than her choices. It is these evaluative commitments that matter to us in moral discourse, as evinced in cases where there has been no explicit choice (instances of forgetting, neglecting, spontaneously reacting).

### **Contrasting the Views**

It was noted at the outset of this chapter that any theory of responsibility will depend on a corresponding conception of agency. Though the rational relations view arguably accounts more fully for the role of responsibility in our moral practices, we should consider whether its corresponding concept of a moral agent is as compelling. And indeed, those suspicious of an

account of responsibility that would indict an agent for an involuntarily held attitude will be quick to question the assumptions regarding our agency that support such ascriptions of responsibility. Therefore, let us consider Scanlon's view in contrast to those underlying the various volitional views.

The prior choice variety of the volitional view held that an agent's responsibility results from their informed choice where they have options. One could do any of two or more actions, and is cognizant of the likely consequences of each action. The ensuing state of affairs is ascribed to the agent for responsibility purposes just because the agent could have chosen otherwise and knew the likely consequences of their action. Yet, it would seem that this account must reduce to something more closely approximating Scanlon's view. Adams (1985, 12-13) gives the example of ingratitude to illustrate this idea. One might find oneself in an estranged relationship with an individual who greatly benefitted him, but against whom he only expressed ingratitude and resentment. The prior choice view would hold the agent responsible for this regrettable state of affairs, since he could have done otherwise (i.e., been thankful and appreciative) and ought to have known the consequences of ingratitude in such circumstances. But in performing this analysis of the prior choice, it seems we are merely criticizing the evaluative judgement on which the agent decided to be ungrateful. The agent judged that being an ingrate was what he had most reason to do, and this is what is objectionable, for if we consider only the resultant state of affairs then we must contend with various contingencies. Perhaps the benefactor has thick skin, and isn't perturbed by the ingratitude and resentment. Should we say in this alternative case that the ingrate is not responsible for being ungrateful? The reducibility of prior choice analyses to evaluative judgements illustrates superiority of Scanlon's conception of agency against this view of responsibility.

The present identification view was presented as relying on Frankfurt's conception of agency. We are responsible for those desires and volitions which we endorse. Where our will is unfree, as when we form a second-order volition that is not effective in determining our action, we cannot be held responsible as agents, but rather only as actors; it is the former guise that matters morally, and so in such cases moral criticism cannot attain. Comparing this conception of agency to that advanced by Scanlon, a fundamental distinction is evident. Recall that Scanlon takes an agent to be conscious, rational, and embodied. This last criterion supports his view that the desires that assail us are a consequence of our physical constitution. And while our physiological motivations (e.g., hunger, thirst) are not subject to moral appraisal, as has been noted what we take notice of and what occurs to us is of moral interest. As embodied agents, we can account for this since the operation of our rational faculties is realized through our physical being. Thus, the considerations that occur to us are very much representative of who we are.

In contrast, the first-order desires that assail us in Frankfurt's view bear no immediate relation to our agency. It is Frankfurt's opinion that while many other animals have first-order desires, and may even possess deliberative abilities, it is the distinguishing capability of persons (read: moral agents) that they are able to form second-order desires. But how is this move from first- to second-order desires achieved? Throughout Frankfurt's early treatments of agency, there exists a conceptual linkage between the active and passive with the internal and external. Whereas we are passive with regard to our first-order desires which are conceived of as external to the agent, we are in contrast active in our second-order desires and volitions which are internally generated by the agent. Thus, for Frankfurt, as with the agentic assumptions of all volitional views, our agency is necessarily identified with our conscious states.

This conjunction of agency and activity has been pressed by Frankfurt's critics. Richard Moran insightfully notes that we usually identify with intentional states over brute sensations, insofar as the former are justified by some rational considerations. One is moved by a desire to eat not simply because one is hungry (a brute sensation), but because it is reasonable to eat something, perhaps since it is time for a meal or to silence distracting hunger pangs. Indeed, were the brute sensation of hunger alone usually motivationally efficacious, then we should expect gluttony to be a normal behaviour. Instead, agents are patently able to rationally situate the desires that assail them, and are usually moved by those that seem most reasonable to them. Consequently, it seems we identify with what we have most reason to in light of norms of rationality. But Moran goes further, to even suggest that our base desires and attitudes are responsive to norms in virtue of the value appraisals necessarily ascribed to them by our rational faculties prior to our conscious awareness. He observes, "In my consciousness of this feeling I like, I have to know my way around, I have to know what there is to enjoy in it, which aspects are central to its pleasure and which are accidental to it, which to attend to and seek more of, and which are indifferent to the pleasure. I have to know how to orient myself toward this feeling in order to enjoy it" (Moran 2002, 211). Thus, our consciousness is shown to be dependent upon the antecedent judgements of our rational faculties.

Considered in light of Moran's criticism, a fundamental flaw in Frankfurt's conception of agency is evident. While holding that agency adheres solely to the active, the conscious, and the internal, Frankfurt fails to recognize the role played by our passive rational faculties. The points raised by Moran show that consciousness alone cannot constitute agency, for it lacks the appropriate orientation to the objects of its awareness. For Frankfurt's second-order desires and volitions to have any credibility, some license for passive mental processes must be granted, in

order to prepare the ground for our active mental experience. This license is implicit in Scanlon's conception, where we might distinguish between an agent's conscious and unconscious states, but nevertheless recognize that their constituents always rest upon the rational faculties.

Future alterability views only hold us responsible for attitudes that are subject to potential future revision. Like Frankfurt's conception of agency, that at play in the future alterability view presupposes that the base motivations which assail us are not chosen or reflective of our agency, but, in contrast, assumes that we can nevertheless revise such attitudes, beliefs, and emotions so that we are morally better persons in the future. The idea of such revision tracks the idea present in Scanlon's conception that our judgement-sensitive attitudes are necessarily responsive to the evaluations of our rational faculties. When we decided that a certain consideration does not count as a reason, it ought to desist from occurring to us in our deliberations in future similar circumstances. Again, this process of our rational faculties depends on the smooth operation of a coherent psychology. For any attitude that is unresponsive to the affective efforts of our reason, such must be considered as beyond the agent's control. And so, the future alterability view and the rational relations view will reach identical judgements of responsibility where attitudes resist revision; however, the rational relations view, in virtue of Scanlon's account of agency, also addresses the attitudes immediately present in an agent, thus expanding the scope of moral assessment.

Whereas all volitional views espoused some blatant version of the Control Principle, the rational relations view requires only the control that normally obtains in a psychologically coherent agent's rational connections between her evaluations and resultant mental states. Smith sought to illustrate the boundaries of these connections with her arachnophobe example, but it can perhaps be drawn out more starkly in considering the case of an acrophobe. Seized by a fear of

heights, acrophobes are often rendered incapable of action by their terror. But what is perhaps most notable is that for the acrophobe there is usually no warrant for their fear, in contrast to the arachnophobe, who may harbour unconscious and substantively mistaken beliefs about spiders. But with the acrophobe, there is no argument that can assuage the terror, for it is completely beyond any rational control. Thus, it is a better example of where the rational control that characterizes our agency on both Scanlon's and Smith's view fails; concomitantly, such instances are not ones where responsibility for the attitude in question can be imputed.

It is essential to the conception of agency endorsed by Scanlon and Smith that our evaluative commitments reflect our judgements about value. What is distinctive about this view is that such judgments need not be reached consciously, and thus may seem, *prima facie*, to be involuntary. But this is mistaken, for if the agent introspects her supposedly involuntary evaluative commitments, then she must find them reflective of her judgements, else they must be external to her as in the cases of hypnotism and brain stimulation. This implication is guaranteed by the conception of agency developed by Scanlon and defended in this chapter. The contrasting concepts at play here between the volitional views and the rational relations view can be further illuminated by revisiting the cases of hypnotism and brain stimulation.

The previously presented examples of hypnotism and brain stimulation were used to illustrate circumstances where responsibility cannot be ascribed on either the volitional views or the rational relations view. Such circumstances were deemed to constrain choice on the former view, and were "external" to the agent on the latter view. These differing interpretations point towards a significant conceptual conflict. Actions, attitudes, beliefs, or emotions brought about through hypnosis or brain stimulation must, *prima facie*, be states that an agent is responsible for on the volitional view, since such states are by their nature experienced consciously and consequently

are subject to the agent's control. That an agent barks like a dog whenever he hears someone say "New York" is something which he ought to be able to prevent himself from doing through conscious effort. When we learn that this agent has been hypnotised to act in this way, the assessment is altered because this fact is considered to have coerced the agent's action. It is then said that the agent did not *want* to bark upon hearing "New York", but instead was forced to against his will. This is the excusing argument for any implanted mental state on the volitional views.

The weakness in this excusing argument is that such appeals to coercion and duress can arguably be considered arbitrary. External pressures on an agent affect their choices in myriad ways, and identifying a reasonable demarcation point for moral responsibility seems impossible. Consider a bank teller who is threatened by a thief demanding money. Handing over the money seems excusable on the volitional view since it was done under duress. But perhaps the teller was frustrated with the bank, and would have happily given away their money to anyone who asked; in such a case, the threat of violence would have no coercive effect on the agent's choice. In contrast to cases where the external pressure is an ineffective influence on the agent, there are situations where the circumstances overdetermine what an agent will do. Such situations might be ones of practical necessity or more mundane ones where the balance of reasons counts so heavily in favour of one course of action that it makes sense to say an agent has no choice in what to do. The effect exerted on the agent by these external pressures effectively extinguishes choice in these particular circumstances. Yet, instances of practical necessity are illustrative of our essential selves, and so we would be giving up too much in evading responsibility in these cases on the basis of our choice being constrained. And so, it seems extremely difficult to delineate where external influences on an agent's choice begin to mitigate her responsibility for that choice.

In contrast, the rational relations view took cases of external influence such as hypnosis and brain stimulation to be interferences with the normal functioning of an agent's psychology. Rather than appealing to the coercive effects of this interference, it identified the manner in which the states brought about by the interference were delinked from the agent's evaluative commitments. Responsibility is mitigated, on the rational relations view, where the resultant states are divergent from those that ought to be produced by an agent's judgment-sensitive attitudes. Thus, barking like a dog upon hearing "New York" is not the sort of action we are likely to ascribe to an agent, but, a strong attitude not to smoke, implanted by therapeutic hypnosis performed at the agent's behest to address their repeated backsliding, might bear sufficient similarity to an evaluative commitment previously defeated by akrasia to warrant ascription to the agent for purposes of moral assessment. In contrast to the conscious control over our choices emphasised by the volitional views, the rational relations view instead addresses the root of our practical reasoning, which allows us to identify the extent to which external forces, such as hypnosis and brain stimulation, are in fact alienating us from our true selves.

Talk of conscious choice in volitional accounts and the psychological linkages that obtain amongst our faculties in the rational relations view ultimately reduce to concerns regarding agent causation: how does one affect the actions, attitudes, beliefs, and emotions that characterize one's mental life. Volitional views take conscious awareness as the context in which an agent's choices are causally effective. To be sure, the control that concerns the Control Principle must be causally effective, insofar as this principle aims solely at picking out the aspects of an agent's mental life that can be said to be representative of her *qua* agent. Yet, as has been observed several times already, conscious choice alone is insufficient to alter our evaluations, and these evaluations are ultimately the determinate of what an agent will do. This is demonstrated in situations where we

(consciously) find ourselves to be objectionable. No doubt many a coward has deliberately resolved to be brave, only to have their desire for safety and to avoid danger overwhelm this conscious edict to stand fast. Persistent evaluative judgements, such as the coward's privileging flight over fight, are not necessarily indicative of failures of agency or our rational faculties. Rather, such persistent commitments are simply reflective of the agent's deep self, and these will not change if they are not constitutive of who one truly is.

Whereas volitional views assume conscious control alone will alter our judgement-sensitive attitudes, the rational relations view posits our reason as affecting alterations of our evaluative commitments. Reason, while the causal mechanism of such change, is certainly aided by consciousness in maintaining and modifying our evaluative judgements. As in the previous example, an agent may find himself ashamed of his previous cowardice, and explicitly resolve to be brave. This resolution, experienced consciously, is a reaffirmation of the reasons one has to be brave, and consciously seeks to have our rational faculties reassess the valuations of the considerations at play. Indeed, such a re-evaluation is just what moral criticism seeks to achieve, and highlights why answerability is so central to the rational relations view. For, even if an agent is consciously unaware or resistant to admitting to a judgement-sensitive attitude, it must be the case that any authentic action for which an agent can be ascribed responsibility is one where the agent can be pressed for the considerations taken to justify that action. Even those unalterable evaluative commitments, those which are representative of the agent's deep self, still warrant moral appraisal, despite their insusceptibility to modification or their offensiveness to conscious reflection, so long as the evaluation is efficacious in producing the relevant mental states when weighed against other reasons. They simply are an essential aspect of who the agent fundamentally is.

Thus, the rational relations view constitutes a refinement and extension of our traditional conception of responsibility. It relies on a more nuanced appreciation of agent causation than the explicit volitional views, a move supported by its underlying account of agency. A consequence of this is that we can be held responsible for involuntary actions, attitudes, beliefs, and emotions. However, these objects of assessment are nevertheless unconscious products of our agency, and we remain answerable for them as such.

## Chapter 2

### Necessary Capacities for Moral Responsibility

Against traditional views of responsibility, which sought to ground ascriptions of moral responsibility in the voluntary choices made by agents, it was argued in the preceding chapter that responsibility instead tracks an agent's evaluative judgements, and that these need not be consciously made in order for an agent to be liable to moral criticism for that judgement. This conclusion was the result of argumentation that relied primarily on observations regarding the nature of our agency, particularly that conception favoured by Scanlon wherein our rational faculties play a central role. For Scanlon, in contrast to more traditional thinkers on the matter, our conscious selves are very much "along for the ride", and a great deal of our moral psychology takes place "below the surface". On the rational relations view, while we are rarely wrong about what we value, it remains the case that our conscious selves nevertheless must introspect to confirm what we like and dislike, what we believe, and what we feel; herein, the close connection between our consciousness and our rational faculties is evinced.

I noted at the outset that moral responsibility bears such an important role because it is broadly conceived as a necessary condition for the expression of blame or praise. Offering praise to an agent who does not deserve it by virtue of being responsible for the laudable act strikes us as absurd. Similarly, several ways in which the expression of blame against an agent for an act she is not responsible for offend against our sense of agency have been examined in the preceding discussion. Yet, nothing that has been said thus far on the conditions for ascription of responsibility have gone any further than establishing that an agent is an appropriate subject for

moral appraisal. Such responsibility assessments do not imply any further action, neither blame in the case of a negative assessment nor praise in the case of a positive assessment.

Despite the distinction the rational relations view draws between responsibility and blame, some contemporary philosophers<sup>13</sup> nevertheless maintain that a necessary linkage exists between the two. Since, on their view, ascriptions of moral responsibility open the agent to further expressions of blame, they consider it a necessary condition that the agent be sensitive to moral reasons. When an agent fails to comply with moral requirements, they question why these moral considerations did not figure in the agent's deliberations. If the agent simply ignored or discounted the moral reasons bearing on a situation, then they agree that he is morally responsible for the action; moreover, the objectors further claim that the agent is subject to expression of blame and sanctions in virtue of this moral responsibility judgment. However, if the agent was unable to perceive or process the moral reasons at play, then they claim moral responsibility cannot be imputed by reason of reduced capacity. Since the agent cannot respond to moral considerations, they could not have acted morally<sup>14</sup>, and thus should not be held accountable for the moral failing through the expression of blame or imposition of sanctions.

Consider the case of an agent who is raised in an impoverished and abusive environment. Growing up, this individual was consistently deprived and regularly beaten. Surviving into adulthood, she evinces many peculiarities of character, being quick to anger, lacking a sense of humour, and having difficulty forming relationships, all attributable to her harsh upbringing. When this individual's behaviour is judged by others, they may initially claim she is morally responsible for being such a disreputable person; but, when they learn of her difficult childhood

---

<sup>13</sup> E.g., David Shoemaker, R. Jay Wallace and Susan Wolf.

<sup>14</sup> They could act in accordance with moral requirements while responding to non-moral reasons, but this would merely be accidental.

and the challenges she faced, we might expect this information to mitigate the responsibility she bears for her dispositions. It seems her moral capacities were never properly developed, and as such she is presently unable to act in accordance with moral reasons. This lack of sensitivity to moral considerations, the objectors argue, should either mitigate or excuse the agent from moral responsibility, since such a negative appraisal would unjustly indict her given her unfortunate biography.

The disagreement thus presents itself in cases involving agents who supposedly lack a capacity to respond to moral reasons. And while the foregoing example illustrates one circumstance where responsibility or blameworthiness judgments might be mitigated by circumstances, a more concrete class of agents is needed to illustrate the theoretical contention that animates the objectors to the rational relations view. Children and the mentally handicapped are generally said to be excluded from this class due to their underdeveloped rational faculties. Similarly, agents suffering mental disorders, such as schizophrenia, dementia, or some psychosis are excluded since their rational functioning is so significantly impaired. After sorting through candidate classes of agents, one presents as potentially meeting the objector's criteria: psychopaths. Psychopaths, like normal agents, exhibit a fully developed capacity for practical reasoning. They can form beliefs about the world, determine the means to their ends, and are possessed of an effective will. Yet, psychopaths differ from normal agents in several fundamental ways, principally in their lack of empathy, their self-centeredness, and their disregard for moral considerations. Consequently, psychopaths are often presented as agents who lack a capacity for moral reasoning, whether this is manifested in a failure to perceive moral reasons or an inability to have such reasons figure in their practical reasoning.

In this chapter I will consider whether, against the rational relations view, a further capacity for moral reasoning is required beyond the general capacity of practical reasoning. In doing so, the psychopath, as the locus of disagreement between the two views under examination, will figure in most examples. I shall proceed first by saying some more about psychopaths, clearly defining the type of agent over whom the debate rages. I shall then go on to give a more detailed account of the view advanced by those who require more than the general practical reasoning capacity extant in the rational relations view, along with a critique of each approach. I shall conclude with a discussion of the relationship between moral responsibility and blame.

### **Psychopathy Defined**

Establishing robust criteria for identifying psychopaths has proven extremely difficult. Psychopathy is not a disorder recognized by the American Psychological Association. Moreover, many psychopaths manage to navigate through life without attracting much negative attention, either through guile, luck, or situational factors. And so, pinning down exactly what type of agent objectors to the rational relations view have in mind proves difficult. I will attempt to illustrate the type of individual they conceive by first discussing the general features attributed to psychopaths, then discussing some clinical approaches to psychopathy that attempt to distinguish it from similar disorders and which hint towards its etiological bases.

Psychopaths are jointly described by both personal characteristics and overt behaviours. Such personal characteristics include a lack of empathy, shallow emotional affect, and egocentricity. Their lack of empathy is evinced in that they feel no guilt for harming others, they often manipulate others for their personal gain or amusement, and in some cases they seem to lack any understanding of others' feelings. In addition to having a poor perception of other's emotional concerns, psychopaths also exhibit shallow emotional affect. They do not *feel* emotions as

strongly as normal individuals in usual circumstances. Finally, psychopaths are characterized by their extreme egocentricity. They tend to act only for their own self-interest, and have grand opinions of themselves. The overt behaviours that describe psychopaths include imprudence, a need for excitement, and general anti-social behaviour. Psychopaths often act imprudently, usually being motivated to act on impulses for short-term gratification against their longer-term interests. This imprudence joined with their lack of empathy is what usually brings psychopaths to the attention of law enforcement. This attention may also be a consequence of their need for excitement, which researchers posit may be an attempt to overcome their emotional deficit by acting at extremes to elicit a more pronounced emotional response. The foregoing description, both in personal characteristics and overt actions, should make the final marker of psychopathy no surprise. Psychopaths are given to anti-social behaviour, where this is defined as a broad and persistent pattern of disregard for and violation of the rights of others. Such anti-social behaviour, it should be noted, is not necessarily illegal, though this is often the case.

Given the above description, one might expect that psychopathy be clearly defined as clinical disorder. Yet, this is not the case. Two factors contribute to the exclusion of psychopathy from the catalog of recognized mental disorders. First, psychopathy does not necessarily impair one's social functioning. Most mental disorders are recognized as such in large part because of the disruptive effect they have on an individual's ability to participate in society. But a psychopath, whose personality may drive him to anti-social behaviours, may still evade notice or sanction and evince, on the surface, normal social functioning. Indeed, this point is what makes the psychopath of interest to the present philosophical question. He is able to purport himself as an otherwise normal agent. Moreover, psychopathy is negatively correlated with clinical mental disorders such as depression, anxiety disorders, attention deficit disorders, and schizophrenia, suggesting that

psychopaths are generally otherwise mentally sound. Second, when identified by mental health practitioners, psychopaths are treated as a sub-class of individuals diagnosed with anti-social personality disorder (ASPD). ASPD is a broadly conceived condition that tracks pervasive anti-social behaviour that has an identified history going back to an individual's childhood. ASPD has an extremely high prevalence in certain populations, with some estimates placing the number of inmates in Canadian correctional facilities diagnosable with ASPD at as high as 80%. However, because of its focus on overt socially deviant behaviours rather than internal dispositions, ASPD fails to account for psychopaths who have yet to be identified by their actions. Thus, researchers have sought to distinguish psychopathy from ASPD and other disorders which require some social impairment for diagnosis.

Two observations suggest a distinction between psychopaths and individuals with ASPD. The first observation is in the different manners in which individuals engage in aggressive behaviour. Most people employ reactive aggression, which is deployed in response to a perceived threat or frustration. Psychopaths, while exhibiting a high degree of reactive aggression, also employ instrumental aggression, which they deploy as a means to achieve their particular goals. The second observation concerns the differing etiological bases of psychopathy and ASPD. Individuals with ASPD often come from backgrounds that fail to adequately socialize them, and so environmental influences are conceived to play a large part in the etiology of ASPD<sup>15</sup>. In contrast, environmental factors seem to play no role in the development of psychopathy. Instances where psychopaths exhibit poor socialization are usually attributable to the resistance caused by their psychopathic tendencies (lack of empathy, impulsivity, shallow affect) to social inculcation. Present research suggests that psychopaths are, in fact, "born this way".

---

<sup>15</sup> N.B., a large role for genetic factors is also assumed.

The impression the above description of psychopaths gives us is one of agents to whom moral considerations are alien. They lack empathy with others, they are completely self-centered, and they view aggression as an acceptable instrumental means to achieving their ends. Their overt social behaviour is often immoral, as they evince little regard for the rights and interests of others. Importantly, it seems that psychopaths possess these dispositions from birth – they are not a product of harsh upbringings or extremely impoverished socio-economic circumstances. It might seem then that psychopaths lack the capacity to act morally.

It is this suggestion which motivates philosophers who disagree with the rational relations view of responsibility. As was noted in the preceding chapter, a central concern of responsibility theories is the type of control an agent has over what she authors. The description given of psychopaths might suggest that since they appear to lack a capacity to respond to moral considerations they consequently cannot be said to control their actions in morally relevant circumstances. Lacking such control, they cannot have responsibility for such moral transgressions imputed against them.

Several philosophers have advanced arguments for the necessity of a capacity for moral reasoning if moral responsibility is to be ascribed to an agent. Susan Wolf argues that this capacity is necessary given the special force that judgements of moral responsibility carry. David Shoemaker, responding directly to Angela Smith, argues against the rational relations view and contends that those who lack a capacity to grasp particular moral reasons should escape full moral responsibility since it would be unfair to sanction such agents. Let us examine each view in turn.

### **The Special Force of Moral Appraisal**

Susan Wolf has advocated for an account of responsibility that tracks an agent's reasons, much like what is proposed by the rational relations view. However, Wolf endorses two further

considerations that divorce her from this view. The first concerns the requisite appreciation of reasons for moral responsibility, and the second concerns the special force she takes judgments of responsibility to have. For Wolf, lacking the former makes imposition of the latter unfair, and this speaks directly to the case of the psychopath.

Wolf believes that the fundamental precondition for ascriptions of responsibility is “that the agent is able to do the right thing for the right reasons, that is, that the agent is able to act in accordance with the True and the Good” (1990, 87). Implicit in this assertion are two component claims. The agent must be able to do the right thing. This implies that the agent must have an effective will. When she judges that she has sufficient reason to perform some action, she does so barring external interference. Thus, this capacity is simply the responsiveness of the will to norms of rationality. Further, the agent must be able to act in accordance with the True and the Good. This claim requires that the agent be able to make accurate evaluations. It is imperative that the agent is able to form true beliefs rather than false ones, and to value things she has reason to value rather than those which are to her detriment. Thus, this capacity is an agent’s epistemic competence with regard to value judgements. Jointly, the capabilities of our will to conform to norms of rationality with the epistemic competence to form accurate judgements about value are what enable us to be responsible agents.

Moral failings on this view are thus instances of either akratic action or evaluative failures. An agent might be possessed of accurate evaluations and decide upon a moral course of action, but through some weakness of will ultimately fails to pursue that end. If the agent has the ability for her will to act upon the dictates of her reason, then she will be responsible for such failures;

however, if she lacks this ability, even circumstantially<sup>16</sup>, then she must be excused from responsibility for the action. As Wolf observes, “the agents in question will be culpable only if they could have acted on their knowledge” (1990, 89). Evaluative failures could be evinced either in substantive errors or epistemic negligence. Such substantive errors simply involve being mistaken about the True and the Good. An agent might (mistakenly) judge that being on time for an appointment is more valuable than checking on an individual lying unconscious in the alley. Cases of epistemic negligence occur when an agent fails to perform due diligence in the evaluations that inform their actions. Failures to take notice, instances of thoughtlessness, and even inconsiderate behaviour fall into this category. Again, moral responsibility for these evaluative failures can only be imputed where the agent possesses the ability to form the correct evaluation. “These failures are culpable only if the agents in question did not have to fail; if the agents, in other words, could have raised the right questions and could have come to the right conclusions, if they could have had the thoughts that they did not actually have” (Wolf 1990, 88).

While the problem of akratic action is one common to all accounts of rational self-governance, the further qualification that an agent need possess a capacity to recognize the True and the Good in order to be morally responsible sets Wolf’s view apart. She puts it explicitly, that “an agent cannot be free and responsible unless she can sufficiently see and appreciate the world (or the relevant portion of it) for what it is” (Wolf 1990, 117). The concern seems reasonable enough, for if we cannot form correct evaluations then our lives will go well only by accident. Beyond this, Wolf is motivated to require her qualifier because of the special force she takes moral responsibility judgements to have.

---

<sup>16</sup> Such circumstances might include those of hypnosis or brain stimulation discussed in the preceding chapter.

Wolf draws a contrast between that for which we are superficially responsible and that for which we are responsible in deep sense. Where I am superficially responsible for hurting you when someone pushes me into you, I am not responsible in the deep sense Wolf has in mind. While it is true that my crashing into you was the cause of your injury, it is not true that the action reflects anything about my will, since my will was not the cause of the action. However, if I attempt to kill you and fail, I superficially avoid responsibility for your murder but bear moral responsibility in the deep sense. Here, regardless of the outcome of my action, the quality of my will in so acting is that which is appraised. In such cases, Wolf contends that holding an individual responsible goes beyond identifying the causal part played by their agency in the event. Such judgements extend beyond the discreet event being moral or immoral, and instead speak to the moral character of the agent herself. These judgements against an agent's moral character are serious, in that they bear a special force of condemnation. Moreover, Wolf takes such ascriptions of responsibility in this deep sense to be linked with explicit blame and sanctions, for in judging an agent morally responsible "[w]e are regarding her as a fit subject for credit or discredit on the basis of the role she plays"(Wolf 1990, 41). Since moral responsibility places one in such moral peril, Wolf thus requires that "in addition to the requirement that the agent be able to govern her behavior by her reasons ... an agent must be in a position that allows her reasons to be governed by what reasons there are"(Wolf 1990, 117) in order to be a morally responsible agent.

To illustrate her concern, Wolf gives us the example of Jojo. The son of Jo the First, an evil dictator ruling over an undeveloped and isolated country, Jojo is favoured by his father's attention and comes to regard him as his role-model. Jojo is given an education and also accompanies his father as he goes about his ruthless rule. Consequently, Jojo grows up with a robust knowledge of

the world in virtue of his tutoring and with values he adopts from observing his father. As an adult, Jojo performs horrible acts, consigning citizens unjustly to imprisonment, torture, and even death. Jojo feels no remorse for these deeds, and he commits them of his own volition; indeed, he sincerely avows his actions as being valuable. Reflecting on his life, Jojo affirms that he is happy with himself.

Wolf intends Jojo's case to illustrate the excusing power of moral ignorance on our responsibility judgements. Jojo is cognitively competent, knowing about the world from his education and demonstrating a capacity for reasoning, but is morally incompetent from growing up in a morally impoverished environment. He cannot act morally because he has never been exposed to the relevant values. Hence, we cannot on Wolf's view hold Jojo accountable for his inhumane deeds, since he lacks the capacity to understand what morality requires of him. And, given the force Wolf takes responsibility judgements to have, finding Jojo responsible would unfairly indemnify his moral character, as he had no control over his upbringing and circumstances which produced his morally bankrupt values. Lastly, as Wolf conceives of such deep responsibility as being linked with deep blame, judging Jojo responsible would imply blame and sanctions that he could not have avoided, violating the Control Principle.

How do psychopaths fair on the view being advanced by Wolf? Given our definition, it would seem that psychopaths will escape moral responsibility for their immoral action in one of several ways. It might be argued that psychopaths fail to meet Wolf's first precondition for moral responsibility, possession of a will responsive to the dictates of rational normativity. For while these individuals often act in accordance with reason, their lives are punctuated by many instances of seemingly akratic action, wherein they fail to do what they have compelling reason to do, and instead pursue some short-term immediate gratification. How much like a smoker,

attempting to quit the habit yet still lighting up “one last cigarette”, is the psychopath, perhaps working diligently but drawn to perpetrate some mischief for a moment’s excitement. To be sure, such impulsivity and need for excitement are among the defining characteristics of psychopathy. But whereas we take the smoker to have the ability to conform his will to the dictates of his reason, and certainly many smokers are successful in quitting the habit, it is claimed that psychopaths are simply constituted such that they cannot resist the lure of certain seemingly akratic actions. In such circumstances, it would seem, the will of the psychopath is ineffective at fulfilling the conclusions of reason. Consequently, on Wolf’s view, this incapacity would make ascriptions of moral responsibility for moral failing in such cases unfair, since the psychopath could not have acted otherwise.

This view might be resisted, however, in favour of Wolf’s second precondition. For whereas the smoker usually regrets his akratic action, such regret is a foreign concept to the psychopath. When they act on their impulses against their long-term interests, they do not feel alienated from themselves by such acts. This suggests that their imprudence and excitement-seeking are not the result of failures of their will, but instead are consequences of their value judgements. The psychopath pursues mischief over some fruitful labour, not because of a malfunction of the will, but because he judges the excitement and fun of mischief to be more compelling than work. Where the actions favour relatively innocent acts such as mischief over hard work, the evaluative failure seems benign enough. But the psychopath will exhibit far more serious evaluative failure in their lack of sensitivity to moral values. Much as such individuals are defined by their imprudence and impulsivity, so too is a lack of empathy and lack of regard for others a hallmark of their character. Thus, when deliberating on what to do, considerations concerning others do not make it onto the scene for the psychopath, presumably because such considerations are not

deemed valuable by the psychopathic agent. But, as Wolf notes, agents are only culpable where they could have acted properly if they had the right evaluations. Given that psychopaths evidently cannot form such moral evaluations, they cannot be held morally responsible for actions in which this failing is evinced. Moreover, unlike Jojo, who could have developed normal moral capacities had he been raised in a different circumstances, psychopaths are apparently born with their incapacity.

Finally, on Wolf's view the special force that judgements of moral responsibility carry would make it unfair to ascribe such judgements against psychopaths, who presumably fail to meet the necessary preconditions for such a judgement to be leveled against them. The deep responsibility such judgements impute, and the corresponding deep blame and sanctions, are seemingly misplaced if the psychopath lacks the capacity to respond to moral reasons. If such judgements go, as Wolf claims, beyond the discreet event to the moral character of the agent himself, then these judgements would indict psychopaths for the inevitable. And, as Wolf conceives of blame as inextricably linked with responsibility judgements, the morally responsible psychopath thus finds himself liable to sanction due to unavoidable consequences of his constitution. Hence, given the severity and seriousness of our responsibility judgements, Wolf's view suggests we should not hold morally responsible agents who lack a capacity for moral reasoning.

Wolf's argument excusing psychopaths from moral responsibility can be attacked at two points. She claims that the special force of moral responsibility judgements is unfair where such judgements are unavoidable. Presumably, the behaviour of psychopaths is unavoidable. Therefore, Wolf concludes, moral responsibility judgements of psychopaths are unfair, given the special force of such judgements and the unavoidability of the behaviour. We can question

whether the behaviour was truly unavoidable, and we can question whether moral responsibility judgements truly have the special force Wolf ascribes to them.

Is the behaviour of psychopaths truly unavoidable? Wolf gave two criteria as comprising the preconditions for moral responsibility on her view: an ability to act rationally and an ability to make accurate evaluative judgements. Our foregoing discussion recognized that the former, while an arguable point for psychopaths, runs afoul of most issues regarding akratic action, and ultimately is unnecessary given the applicability of the latter criteria to their case. Whether or not psychopaths are able to perceive moral reasons turns on both our preceding definition of psychopaths and what it is to “perceive” moral reasons. It was asserted that psychopaths possess both a general practical reasoning ability and a general lack of empathy that evinces itself in disregard for the rights and well-being of others. Hence, psychopaths are able to generally make evaluative judgements about the contents of the world and then make and execute plans on the basis of these appraisals. They are able to perceive most reasons, but is there something special about moral reasons that make them imperceptible to psychopaths? It seems that moral reasons are simply considerations that an action would be unjustifiably injurious to the interest of another agent. To grasp such a consideration requires not empathy but a general understanding of the world, the consequence of actions, and basic social norms. Clearly, psychopaths possess an understanding of these things, as demonstrated by their competence in non-moral practical domains, including interpersonal ones. A lack of empathy does not blind one to moral reasons, though perhaps one could argue that it robs such reasons of their motivational power. Nevertheless, if a psychopath is able to perceive reasons in light of his evaluative commitments, then the failure of moral reasons to figure in his actions is simply that he evaluates them as irrelevant. Telling a psychopath “Your plan puts me at risk!” will only elicit (an honest) reply of

“That does not matter to me”: he understands the propositional content of your statement, but it does not accord with his conception of the Good. Rather than claiming that psychopaths lack an ability to perceive moral reasons, perhaps we ought to instead claim that they simply do not value moral reasons.

What of the special force that moral responsibility judgements allegedly bear? Wolf worries that such judgements transcend the event and strike against the agent’s character. This worry is absolutely correct, and is a consequence of the reducibility of responsibility judgements. When we make an answerability demand of an agent, we ask what reasons they took to justify their action. Implicit in their response is a judgement that the decisive reason or set of reasons outweighed the set of other considerations at play in that particular circumstance. In other words, the agent affirms her valuation of the decisive set over the outweighed set. And as argued in the preceding chapter, these valuations are the constitutive elements of our true selves. So, it is true that negative moral responsibility judgements will speak against an agent’s character. But, on the rational relations view, this is merely a descriptive judgement. The real worry that seems to motivate Wolf is that such judgements are essentially normative. For Wolf, judging that an individual is of poor moral character in virtue of a negative moral responsibility judgment requires that one express blame or in some way sanctions the agent. In considering moral responsibility judgements to bear such normative import, Wolf can be seen as subscribing to the Desert Thesis, the principle that it is morally better that one suffer some loss in consequence of committing a moral wrong. She speaks of deep responsibility being linked with deep blame and sanctions, yet it is not obvious why this should be so. The tendency to conflate moral responsibility and blame is often asserted without any argument for their coextension being given. The rational relations view takes moral responsibility appraisals to be merely descriptive

judgements that lack any further inherent prescriptive element with regard to blame. Such descriptive judgements merely tell us what an agent is like, not what we ought to do about what they are like.

### **The Necessity of Blame**

According to David Shoemaker, an agent's capacities, particularly their capacity to respond to reasons, are essential to appraisal and consequently to blame. Whereas Wolf claimed that a capacity to know and act on the True and the Good was required to be morally responsible, Shoemaker gives a more sophisticated account that turns on the manner in which moral considerations figure in an agent's deliberations. Shoemaker believes that certain incapacities, particularly the incapacity to understand certain reasons, render blame inappropriate. To see why, further consideration of our practical reasoning practices is necessary. First, a brief discussion of the Scanlonian theory of blame and Shoemaker's particular three-fold sense of responsibility must be presented, before turning to Shoemaker's substantive argument.

Shoemaker advances his argument using elements of Scanlon's theory of blame. Blame is the reaction we generally experience when we learn that someone's attitude toward us imparts a meaning that negatively diverges from what we expect given our relationship with the individual. Scanlon offers a definition:

*"To claim that a person is blameworthy for an action is to claim that the action shows something about the agent's attitudes toward others that impairs the relations that others can have with him or her. To blame a person is to judge him or her to be blameworthy and to take your relationship with him or her to be modified in a way that this judgment of impaired relations holds to be appropriate."(Scanlon 2008, 128)*

The type of relationship that obtains between the two is crucial as it establishes the norms expected of the parties. For example, if the relationship is one of friendship then one would expect the parties to it to be well-disposed towards each other and willing to help each other in

the achievement of their various projects. The wrongdoer and the victim might be in such a relationship of friendship, or they might be strangers to each other. In the latter case, Scanlon maintains that a basic moral relationship nevertheless obtains between the two which requires that each be inclined to avoid imposing harms upon the other, offer aid when one can do so at little cost, and generally respect each other as fellow rational beings. When an agent fails to live up to the normative ideals demanded by a particular relationship, she gives the victim reason to alter his attitudes towards her. In other words, such failures on the part of the wrongdoer result in impairment of the relationship, and depending on the severity of the impairment the relationship may or may not continue. Returning to the example of friendship, when one party fails to keep the confidences of the other party, the betrayed agent has reason to revise his attitudes towards his friend, perhaps resolving to no longer trust her with sensitive information about himself; moreover, if the betrayal was significant enough, it may so impair the friendship as to warrant its dissolution. It is important to note here that Scanlon's account is not retributivist. That an agent is judged to be blameworthy does not imply that any blame or resentment need necessarily be expressed against her, nor any other further sanction. This may be called for by reasons arising from the standards governing the particular relationship, but the expression or imposition of sanctions is a further issue from the appraisal itself.

Moral responsibility is analysed into three distinct senses by Shoemaker. First, he believes agents can be attributability-responsible for an action, attitude, belief, or emotion. This sense is separate from answerability-responsibility, for the solely attributability-responsible agent is not able to present justificatory reasons for the action, attitude, belief, or emotion in question. Shoemaker points to instances of irrationality and deeply held emotional commitments as examples where we are attributability-responsible in virtue of our evaluative commitments, but

not answerability-responsible due to the failure of our rational faculties to produce justificatory reasons for the resultant mental states. An agent is answerability-responsible in the sense captured by the rational relations view, that is, where the agent can identify the reasons which she takes to justify her actions, attitudes, beliefs, and emotions. Because answerability-responsibility depends on being able to identify the evaluative commitments that warrant resultant mental states, it entails also being attributability-responsible as well. Finally, Shoemaker believes an agent can be held blameworthy in two distinct senses. One is accountability-responsible for actions taken to be justified by reasons that contravene relationship-defining standards. The other type involves holding one answerable, or answerability-responsible, for actions whose meaning fails to meet relationship-potential standards. This latter type need not be publicly expressed (as with private blame) and does not necessarily warrant any reactive attitude or sanction. Rather, holding someone to account on his view “is precisely to sanction that person, whether it be via the expression of a reactive attitude, public shaming, or something more psychologically or physically damaging” (Shoemaker 2011, 623). Consequently, moral responsibility judgements which find an agent accountability-responsible carry a significant retributive element on Shoemaker’s view.

Shoemaker’s concern is borne out in the contrast between the differing senses with which one can “express ill will” when violating demands imparted by relationship-defining standards. In one sense, ill will can be expressed by seeing the reasons given by the demand to be defeated by competing considerations. Thus, one might recognize the legitimacy of a demand not to cheat on one’s partner, yet lend more weight to the reasons counting in favour of doing so. Another sense would involve the demand as being simply uncounted. Using the same example, it might be that the demand not to cheat does not even factor into the agent’s deliberations. In this case, the

reason does not even make it onto the scene, whereas in the former, it is discounted in favour of perceived stronger reasons.

Shoemaker seeks to illustrate the distinction between these two senses in his *Aliens* example. He supposes that we encounter a race of alien beings that are just like us both psychologically and physically. They share our moral sensibilities, but also claim to see additional moral reasons in the world. Specifically, they believe it to be immoral to harm grass, due to their claimed special capacity to understand what it is like to be a blade of grass. We (human beings) simply do not understand what they are on about with regard to the special moral status of grass. Nevertheless, these aliens are beings with whom we can have moral relationships: we can form friendships with them, make and keep promises, and express blame and praise to each other. Given this preamble, Shoemaker asks us to consider what follows when a human being fails to respect the moral status of grass. Since human beings generally see no (undefeatable) reason to avoid harming grass, the consideration generally does not play any role in their practical reasoning. Yet, for the alien, their perceived special moral status of grass introduces a powerful demand on others to avoid harming it, and failures to do so undermine the moral relationship they can consequently have. Hence, an alien may feel compelled to sanction a human being for harming grass, this being a violation of the moral relationship-defining standard which dictates one must respect the moral status of grass.

Such a reaction, Shoemaker claims, would be a mistake. Human beings, in the example, are incapable of seeing grass as having moral status. Given the dictate “ought implies can”, human beings could not enter into relationships governed by norms to which they cannot adhere. Shoemaker points to the similarity here with our orientation towards children. We cannot enter into certain types of relationships with children, simply because children lack the capacity to understand particular reasons constituting the norms which define these relations. So, just as it

would be inappropriate to sanction children who, through their incapacity to understand the considerations at play, violate a foundational relationship norm, so too would it be inappropriate for the aliens to sanction human beings, who are incapable of understanding grass's moral status, for trampling so many blades. Because the human being cannot see the supposed value of grass, while he fails to respect it, he does not actively disrespect it.

While expressing blame or imposing sanctions would not be warranted, Shoemaker concedes that the alien would be justified in modifying its attitudes towards the human who tramples grass. This human's actions, and more specifically the disregard of the supposedly moral reasons given by grass's purported worth, have significance for the type of relation the alien can have with this individual. Subsequently, the relationship will be impaired by the fact that the human does not recognize the reasons to avoid harming grass. This may result in the alien judging the human to be insensitive or callous, and perhaps to avoid interacting with this individual in the future. But, if the alien knows that the human lacks the capacity to understand that it is wrong to harm grass, then it cannot hold that person accountable. Such accountability is not possible since one party to the relationship is incapable of seeing the reasons constituting the normative standard of the relationship as reason giving. Relationships are founded on the mutual demands that the parties can make upon each other, but the incapacity under discussion makes impossible this mutuality. Moreover, Shoemaker observes that "such an incapacity undermines the possibility of my expressing ill will in the sense warranting accountability-blame, namely, active disregard" (Shoemaker 2011, 627).

Here, Shoemaker's deviation from Wolf's view is most evident. Whereas Wolf contended that general moral ignorance excused one from moral responsibility, Shoemaker's more sophisticated account requires that the moral incapacity not express any ill will. The distinction is clear in

considering Wolf's Jojo example against Shoemaker's theory. Shoemaker, in opposition to Wolf, would find Jojo to be accountability-responsible for his inhumane acts, since Jojo, like any other normal agent, possesses the capacity to understand moral reasons and in performing his heinous deeds does knowingly disrespect others. Whereas Wolf excuses Jojo due to his lack of moral education, Shoemaker refuses to excuse him just in virtue of his (undeveloped) capacity to understand the moral reasons which apply to him and his relationship with others. Jojo's citizens are not Aliens, making unintelligible moral demands of him, but are fellow agents whose worth, he has been taught, is negligible.

Shoemaker directly addresses how his view applies to psychopaths. His analysis of the moral responsibility that psychopathic agents bear results in a multivariate answer. He concedes that psychopaths are responsible in the senses identified by the rational relations view, senses that Shoemaker wishes to distinguish. Psychopaths are attributively responsible in that their actions and attitudes appropriately reflect their evaluative commitments. Such evaluative commitments may be shallow or objectionable, but they nevertheless produce rationally resultant mental states. Thus, Shoemaker believes we are warranted in making attributions of the psychopath, claiming he is cruel, manipulative, or reckless. Answerability demands can also be made of the psychopath, for the evaluative commitments which animate their actions and attitudes are grounded in reasons they take to be justificatory. Pressed for a justification for some reckless action that put others in danger, a psychopath might intelligibly reply that the safety of others does not matter to him and that the excitement generated by the reckless act mattered greatly. As rational agents, psychopaths are just as capable as any of answering the reasons they took to justify their actions, attitudes, beliefs, and emotions.

But Shoemaker does not believe that psychopaths are responsible in what he calls the accountability sense, since being so “is just to be susceptible for being the appropriate target of sanctions ... communicating relationship-defining demands” (Shoemaker 2011, 628). Since psychopaths are incapable of seeing moral demands as reason-giving, they are thus inappropriate targets for accountability. Like the human beings in his *Aliens* example, psychopaths simply lack the capacity to actively disrespect moral demands, and instead can only fail to respect moral demands. And so, while we would be justified in modifying our attitudes and dispositions towards the psychopath (what he calls the Scanlonian sense of blame), Shoemaker believes we are not warranted in imposing sanctions, since such expressions of explicit blame turn on imputations of his accountability-responsibility. In sum, while the psychopath may be responsible in the sense that the evaluative commitments which animate his immoral activity are truly his own, and responsible in that he can give the reasons grounding the evaluations at play in his deliberations, Shoemaker contends that psychopaths are nevertheless immune from explicit blame and sanctions by virtue of their incapacity to be moved by moral considerations.

Shoemaker posits a distinction between answerability and accountability. He sees the two types of response as tracking different agent capacities, specifically the ability to perceive particular considerations as reason giving. In cases of blame, another agent’s attitude might express a relationship impairing meaning in a variety of ways. It might impair the relations we can have with the agent if the meaning expressed implies that the agent does not consider our demands (or the reasons given by our interests) as overriding competing considerations. Alternatively, our relations might be impaired if the meaning implies that particular considerations were not even taken as reasons. Shoemaker contends that in cases of the first sort, accountability can be imputed since the blamed agent took all the relevant reasons into

consideration during deliberation. In the second instance, accountability cannot be imputed, since the blamed agent was unable to properly assess all the considerations bearing on the attitude. However, such an agent would still be answerable for the anemic attitude, and others would be justified in modifying their attitudes in response.

Shoemaker seeks to illustrate this distinction with his *Aliens* example. The aliens perceive additional moral reasons in the world, specifically ones that speak against harming grass. But this example fails to demonstrate the distinction that Shoemaker is seeking to establish. The aliens know something that we do not. Consequently, the difference between them and human beings is an epistemic one. When the human beings in his example “chortle about the aliens’ ridiculous moral beliefs” (Shoemaker 2011, 626), they do so not because the consideration that grass has moral status leaves them cold, but rather because they judge the evidence given by the aliens purporting the moral status of grass as insufficient to warrant their belief. Such an analysis implies that the considerations constituting the reason not to harm grass (as expressed by the aliens) do make it into the human being’s deliberation, but the deliberation is a preceding one of what to believe rather than the practical deliberation of what to do.

Regardless of the efficacy of Shoemaker’s example in establishing his distinction, it is not the case that reasons being outweighed as opposed to being uncounted has any impact on moral responsibility judgements. The rational relations view posits that our evaluative commitments strongly influence our perception of reasons. If we value something, we are more likely to take notice of considerations that relate to this object of value in our deliberations; conversely, if we do not value something, then we are unlikely to have considerations relating to it figure in our practical reasoning. So, whether one explicitly considers a reason and subsequently discounts it in favour of other stronger reasons, or whether the reason is never seized upon for consideration, in

both cases the reason is one that does not hold any decisive value for the agent. Thus, in both cases we are justified in claiming that the agent does take a particular consideration to count.

### **Moral Responsibility and Blame**

Responsibility judgements made on the rational relations view are ultimately descriptions of what an agent is like. These descriptions are of the evaluative commitments held by an agent which produced the practical reasoning that resulted in a certain action, attitude, belief, or emotion. So long as the evaluative commitments are rationally linked through an agent's stable and coherent psychology to the resultant mental states, ascriptions of responsibility for these states is justified. Psychopaths, despite their personal deficits, do in fact possess the stable and coherent psychology that characterizes all other practical reasoners. On the rational relations view of responsibility, the psychopath is certainly morally responsible for his attitudes and actions without qualification.

The preceding critique of Wolf's argument suggested that it is not the case that psychopaths are unable to perceive moral reasons, but rather that psychopaths simply do not value such considerations. According to the rational relations view, this fact is what is at stake in ascriptions of moral responsibility. That considerations about the rights and well-being of others are not counted among the evaluative commitments of a psychopath is precisely the descriptive judgement in which a negative appraisal of moral responsibility consists. And, as Wolf worries, this judgement does speak against the psychopath's moral character, against his "true self". But, such descriptive judgements, despite their presumed normative import, are really not much different from other descriptive judgements we might make, such as "He is tall" or "He has an IQ of 100". Moral responsibility judgements are descriptive of an agent's moral characteristics in the same way the preceding judgments described physical and intellectual characteristics. Moral

responsibility judgements differ from these physical and intellectual descriptions in telling us what an agent is like, rather than what a body or a psychology is like. Though traditional views of responsibility embraced a contrary intuition, the descriptive judgement, “He acted immorally”, has as much normative consequence as “He has blonde hair”.

As conceded by Shoemaker, the psychopath’s evaluative judgements correspond to his overt behaviour, and when pressed for the considerations that informed his deliberation, the psychopath will be able to produce reasons he took to be justificatory. So, in the sense required for both Shoemaker’s attributability and answerability responsibility and Smith’s rational relations responsibility, the psychopath is an appropriate subject for our moral appraisal. But Shoemaker presses for the further distinction of accountability, and holds that psychopaths do not warrant the treatment accountability implies in virtue of their incapacity to respond to moral reasons. Here again, Shoemaker conflates expression with appraisal. The psychopath is morally repugnant purely in virtue of his appropriately held evaluative commitments and the moral appraisal of him as such is appropriate in virtue of the rational relations between his attitudes and these commitments. In response to such appraisals, other agents are clearly justified in modifying their attitudes and dispositions towards the psychopath; however, the rational relations view does not make any prescriptive suggestion, its judgements being purely descriptive.

While Wolf and Shoemaker both argue in different ways for requiring a capacity for moral reasoning, both are ultimately motivated by the worry that moral responsibility judgements necessarily imply blame. For Wolf, the judgment itself carries special force that is constitutive of blame, whereas for Shoemaker, the judgement in certain circumstances implies an obligation to hold accountable the offending agent. Both assumptions are mistaken on the rational relations

view, as this theory establishes a distinct break between moral responsibility and our practices of praise and blame.

How credible is this distinction between moral responsibility and blame? Indeed, it was noted at the outset of the opening chapter that moral responsibility is taken to be the condition that warrants blame and praise. Therefore, in the face of our traditional understanding of the relation between moral responsibility and blame or praise, more must be said if the distinction the rational relations view wishes to establish is to be accepted. I believe the distinction can be supported by a further analysis of the two assumptions made by Wolf and Shoemaker. Once this is complete, I hope to show that psychopaths (and, indeed, any sufficiently rational agent) can bear full moral responsibility for their actions; however, the question of blame or praise is a further fact from this responsibility judgement.

Moral responsibility judgements are considered to carry a special force or weight, such that philosophers like Wolf feel that even a negative moral appraisal is blame-like in its effect upon an agent. This is so due to the “deep” nature of the predications at play. When we say an agent acted without regard for others, this speaks to their moral character, to their “true self”. Such a predication, “does not have appropriate regard for others”, is to be avoided, since generally we accept that we normatively are required to have appropriate regard for our fellow agents. The attribution of such a predicate is equivalent to being called a bad person. And this is where the special force of moral responsibility judgements is supposedly felt, for no one should want to be identified as a bad person. Indeed, this proposition has a long history<sup>17</sup>, and it seems on the account of agency that has been advanced in this paper that any such contradiction would be

---

<sup>17</sup> See Plato's *Meno*, 77b-78. Therein, Socrates remarks “It is clear, then, that those who do not know things to be bad do not want what is bad. What they want are things they think are good, that are in fact bad. It follows that those who have no knowledge about these things and believe them to be good clearly want good things”.

paradoxical: since any agent acts on considerations she takes to constitute value, and if it is revealed to her that her reasons were in fact bad, then we should expect her to acknowledge the fact and revise her evaluative commitments accordingly. A failure to revise one's evaluations once one recognizes the disvalue of a particular consideration would indicate a serious malfunction of one's rational faculties.

So, descriptive judgements of moral responsibility generally have some efficacy in motivating the appraised agent in re-examining her evaluative commitments. But is this in fact constitutive of a kind of blame (or praise)? It does not appear so. First, it assumes that the appraised agent shares the normative standards of the appraising agent. In the case of the psychopath, this is clearly not so. Pressed with the judgement that he acted without appropriate regard for others and is consequently a bad person, the psychopath will simply assert that such considerations do not concern him and thus will not revise his evaluative commitments in light of such moral criticism. Given an objective morality, the psychopath's lack of regard is a genuinely bad fact about him; nevertheless, this fact will fail to have the revisionary efficacy it does for normal agents. Second, and following from the first point, simple judgments of moral responsibility do not necessarily carry the weight that special force adherents claim. The negative moral appraisal of the psychopath will not result in any lost sleep. Even normal agents who act immorally sometimes exhibit a lack of concern for the negative appraisals of others<sup>18</sup>. In both these cases, the lack of concern for others' moral opinions robs moral appraisal of its supposed special force. More than simple moral responsibility judgements are required to motivate such agents to revise their evaluations, and this motivation is provided by the distinct expression of blame. Moral responsibility judgements alone lack the special force that in fact is a component of blame.

---

<sup>18</sup> Here, I have in mind not any explicit expressions of blame, but rather a knowledge that others know one acted immorally.

Shoemaker believes that moral responsibility implies holding agents responsible for their actions. The accountability that he suggests we bear will depend on the normative standards of the relationship in which we stand to the appraised agent. He claims, “to hold someone answerability-responsible is to measure the reasons he took to justify his actions or attitudes against the reasons grounded in the standards of the relationship. Where there is a gap between his reasons and the basic normative demands defining the relationship itself, one response is fitting; where there is a gap between his reasons and the standards defining the potential of the relationship, another response is fitting”(Shoemaker 2011, 622). Shoemaker believes violations of the former gap warrant holding an agent to account, which is “precisely to sanction that person, whether it be via the expression of a reactive attitude, public shaming, or something more psychologically or physically damaging”(2011, 623). Thus, making an appraisal of moral responsibility against an agent, where that agent’s reasons violate relationship-defining normative standards, necessarily requires expression of blame.

We can certainly agree with Shoemaker that the violation of relationship-defining norms usually gives rise to feelings of blame. But we can also question the nature of these reactions. While Shoemaker contends that this accountability is inherent to moral responsibility judgements, his own words betray a distinction. In the above quote, he notes that in evaluating an agent’s accountability, we measure the responsibility judgement (“the reasons he took to justify his actions or attitudes”) against the normative standards of the relationship we stand in to the appraised agent. Indeed, in establishing a foundational conception of blame, Shoemaker points out that establishing blameworthiness is done *in addition* to establishing answerability for an action or attitude<sup>19</sup>. What becomes evident, then, is that judgements of blame, of holding an agent

---

<sup>19</sup> See Shoemaker, 2011, pg. 618

accountable, are further judgements. More specifically, they are judgments that arise from antecedent judgements of moral responsibility. Judging an agent morally responsible for some action may give us reason to respond to that appraisal through a modification of our own attitudes; however, this is a further fact from the responsibility appraisal itself. Moreover, the decision to blame will be informed by various considerations, such as the nature of our relationship with the appraised agent, the severity of the moral transgression, and facts concerning the capacities of the agent appraised. Blame will vary if the appraised agent is a close friend or a stranger, if the act was a white lie or an unjustified assault, if the agent appraised is competent adult or an angst-filled teenager. In one circumstance, blame might call for strict sanctions, in another, perhaps nothing explicit. All such deliberations, such a familiar component of our moral practice, depend on a preceding and distinct judgement of moral responsibility.

Psychopaths, on the foregoing analysis, find themselves in an interesting position. According to the rational relations view, they bear full moral responsibility for their actions. Appeals to some missing capacity to perceive or respond to moral reasons cannot mitigate their liability to negative moral appraisal. The rationale that grounds responsibility judgements of the normal agent who murders is the same that grounds the negative appraisal of the psychopathic killer. The supposed special force of such responsibility judgements is contingent upon the normative orientation of the appraised agent. Moreover, the force such judgements do sometimes bear can hardly be characterized as unfair: individuals who act immorally are bad just in virtue of these acts. Such descriptive judgements, however, do not necessarily imply any blame, which is a further judgement beyond simple moral responsibility. In the case of the psychopath, their incapacity to be moved by moral considerations should inform our expressions of blame for their immoral actions. Unlike most normal agents, who are likely to respond to our reactive attitudes

and overt blaming by re-examining their intrinsic evaluative commitments, psychopaths are simply left cold by such expressions and view the consequences of blame as purely instrumental reasons. When we ask why a psychopath does not act on every immoral impulse, the answer is that the threat of sanctions or subsequent impairment of relations from committing the act outweighs the current motivation for performing the immoral action. Hence, we must recognize that in blaming the psychopath, we cannot hope to force a re-evaluation of his deep commitments; rather, all we can hope to do is give him sufficient instrumental reasons to avoid immoral action.

This chapter addressed the worry that an additional capacity for moral reasoning was required for ascriptions of moral responsibility to be made. As discussed, the worry was motivated by the supposed special force of moral responsibility judgements and the presumed obligation to blame that such judgements entailed. Psychopaths, as generally competent practical reasoners who lack a sensitivity to moral considerations, presented as candidates against whom ascriptions of moral responsibility could not be made: both the special force of such judgements and the obligation to sanction seemed unfair in the face of their inability to act in accordance with moral requirements. I hope to have shown that the worries regarding the special force are unfounded given the contingency of its impact, and that the obligation to blame is specious since the decision to do so is a further judgement from that of moral responsibility. Consequently, we can hold psychopaths, indeed any sufficiently competent agent, fully morally responsible for their actions, attitudes, beliefs, and emotions, insofar as these are appropriately linked to the agent's evaluative commitments through a stable and coherent psychology. No further special reasoning capacities are required beyond the general ones identified by the rational relations view for the imputation of moral responsibility.

## Coda

This work has sought to present the rational relations view as the most accurate account of moral responsibility. Since the rational relations view takes a substantially different approach to moral responsibility than competing views, the onus lies upon it to establish itself. Thus, in defending the rational relations view, a contrast was drawn between both traditional volitional views of responsibility and those who believe a capacity for moral reasoning is a necessary condition for responsibility ascriptions. By illustrating the weaknesses and conceptual confusions which I believe plague these two views, the strength of the rational relations view is evinced through its greater explanatory power and conceptual clarity. And while the main subject throughout has been the appropriate conditions for moral responsibility, several additional and interesting consequences emerge from the discussion.

In comparing the rational relations and volitional views of responsibility, the underlying and fundamental role played by conceptions of agency was highlighted. Indeed, much of the argumentation in the first chapter turned upon the understanding of how an agent affects control over her actions, attitudes, beliefs, and emotions. Against prior choice views, which took conscious unconstrained choice with knowledge of consequences as sufficient control, and present identification views, which took an agent's second-order endorsement of extant first-order desires and volitions as sufficient control, the rational relations view instead grounded responsibility in the control evinced by an agent's rational faculties through the coherent connections that obtained between her evaluative commitments and the resultant actions, attitudes, beliefs, and emotions that are produced in her. It was argued that this Scanlonian conception of agency, which grounds the rational relations view, was better able to account for broad judgements of responsibility, in what agents perceive, in what they neglect, and in how

they spontaneously react. While this discussion served to strengthen the rational relations view against the volitional varieties, it also demonstrated some interesting truths regarding the nature of our agency itself.

Among these interesting truths is the nature of our conscious relation to deep selves. At numerous points it was demonstrated that we are often driven to act by unconsciously held evaluations. The rational relations view sought to argue that we are nevertheless responsible for such actions despite not having consciously chosen them. Yet, further examples can be picked out where our unconscious evaluations not only drive our actions, but in fact determine them even against conscious resistance. Such instances of practical necessity might seem, from the conscious perspective, to usurp one's agency and control over a situation. Yet, on the rational relations view and the Scanlonian conception of agency underlying it, such instances of practical necessity are rightly viewed as most fundamentally expressive of an agent's deep self. Consequently, a schism is implied between our conscious awareness and our deep selves.

The argument against a capacity for moral reasoning as a necessary condition for judgements of responsibility focused on the supposed coextension of moral responsibility and blameworthiness. Whereas Susan Wolf claimed that negative moral responsibility judgements possessed a special force that is tantamount to blame, the rational relations view instead takes such judgements to be purely descriptive, lacking any direct normative import. Such judgements have motivational efficacy upon agents only where the descriptive judgement does not match their evaluative commitments<sup>20</sup>. David Shoemaker advanced a view of moral responsibility that

---

<sup>20</sup> Thus, someone descriptively judged as performing a morally wrong action will be motivated in light of this judgment to reform their behavior only if they are truly committed to being a morally good person. Equivalently, someone descriptively judged to be short will be motivated by this judgment to wear lifts only if they deeply value being tall. In both cases, the normative consequence of the judgment depends on further facts regarding what the agent values.

took blame to be necessarily entailed by certain types of responsibility judgements, but that such judgements were not fair unless the agent possessed the capacity to be moved by particular reasons. Yet, analysis of his arguments demonstrated an implicit distinction between a moral responsibility judgement (Shoemaker's answerability-responsibility) and a blameworthiness judgement (Shoemaker's accountability-responsibility). It was this latter form of responsibility which Shoemaker claimed was unfair to press upon agents lacking the capacity to be moved by moral reasons, and so having shown this to be a separate further judgement from moral responsibility assuages the worry he sought to press. Again, while these arguments bolstered the rational relations view against the challenge brought by proponents of a capacity for moral reasoning, the disengagement of moral responsibility and blameworthiness has interesting ramifications for broader moral theory.

Perhaps the most striking consequence of this disengagement and conception of responsibility judgements as descriptive is the expanded class of individuals who can be subject to moral responsibility appraisals under the rational relations view. For, on the rational relations view all that is required is a coherent psychology wherein an agent's rational faculties produce appropriate mental states in accordance with the agent's evaluative judgements. And while Scanlon himself has explicitly excluded children and those with diminished mental capacities from moral appraisal<sup>21</sup>, it seems such agents must be included if we are to be true to the rational relations view. Insofar as such agents act in accordance with their evaluative commitments brought about by normal functioning rational faculties, then it would seem they must bear moral responsibility when their evaluative judgments demonstrate a moral transgression. That 10-year old Timmy

---

<sup>21</sup> See Scanlon 1998, 280. He identifies three reasons to excuse such agents: they are poor at assessing reasons, they lack an understanding of the consequences of their actions, and they may have undeveloped moral sensibilities.

feels it is justifiable to hit his little sister for his own amusement is a moral fault that can be appropriately attributed to him; however, the descriptive judgement of Timmy as morally responsible for the assault is distinct from judging him to be blameworthy for the assault. Here, the excusing considerations that were previously thought to preclude moral responsibility should instead inform the deliberations on whether to blame.<sup>22</sup>

Taken in sum, the rational relations view presents a novel account of the appropriate conditions for moral appraisal. Yet, the view's novelty also presents several unconventional consequences. According to the rational relations view, we are agents with imperfect knowledge of ourselves. At certain times, such as in situations of practical necessity, we may act in unexpected ways which seem constraining to our conscious selves. In divorcing responsibility from blameworthiness, the rational relations view claims moral responsibility judgements are purely descriptive and thus not directly normative. By stripping these judgements of their normative import, a more consistent application of the view can be made, as against agents traditionally excluded from ascriptions of moral responsibility. Ultimately, the rational relations view tells the truth about agents, truths they may not have previously known about themselves and truths which they may perhaps find uncomfortable, but that nevertheless are fundamental reflections of their deep selves which are essential to their agency.

I hope to have shown the rational relations view to be a superior account of the conditions for moral responsibility. While it diverges in several ways from traditional accounts, these variances are justified by supporting theories of practical reasoning and moral psychology. The rational relations view thus not only provides a more comprehensive and systematically coherent account of moral responsibility, but also establishes better fit with broader moral theory in general.

---

<sup>22</sup> This is analogous to the conclusion regarding the treatment of psychopaths at the end of Chapter 2.

## Bibliography

- Adams, Robert Merrihew. "Involuntary Sins." *The Philosophical Review* 94, no. 1 (1985): pp. 3-31.
- Aristotle. *Nicomachean Ethics*. Indianapolis, Ind.: Hackett Publishing, 1999.
- Frankfurt, Harry G. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy* 68, no. 1 (1971): pp. 5-20.
- Hare, Robert. *Psychopathy: Theory and Research*. New York: Wiley, 1970.
- . *Without Conscience: The Disturbing World of the Psychopaths among Us*. New York: The Guilford Press, 1999.
- Hieronymi, Pamela. "The Force and Fairness of Blame." *Philosophical Perspectives* 18, no. 1 (2004): 115-148.
- Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Peterborough, Ont.: Broadview Press, 2005.
- Moran, Richard. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, N.J.: Princeton University Press, 2001.
- . "Frankfurt on Identification: Ambiguities of Activity in Mental Life." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 189-217. Cambridge, Mass.: MIT Press, 2002.
- Nagel, Thomas. *Mortal Questions*. New York: Cambridge University Press, 1979.
- Plato. *Plato's Meno*. New York: Cambridge University Press, 2006.
- Scanlon, Thomas. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, Mass.: Belknap Press of Harvard University Press, 2008.
- . "Reasons and Passions." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 165-183. Cambridge, Mass.: MIT Press, 2002.
- . *What we Owe to each Other*. Cambridge, Mass.: Belknap Press of Harvard University Press, 1998.
- Shoemaker, David. "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility." *Ethics* 121, no. 3 (April, 2011): pp. 602-632.

Smith, Angela M. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115, no. 2 (January, 2005): pp. 236-271.

Strawson, P. F. *Freedom and Resentment, and Other Essays*. London: Methuen, 1974.

Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press, 1994.

Williams, Bernard Arthur Owen. *Moral Luck: Philosophical Papers, 1973-1980*. New York: Cambridge University Press, 1981.

Wolf, Susan. "Sanity and the Metaphysics of Responsibility." In *Responsibility, Character, and the Emotions : New Essays in Moral Psychology*, edited by Ferdinand David Schoeman, 46-62. New York: Cambridge University Press, 1987.

———. *Freedom within Reason*. New York: Oxford University Press, 1990.