

Flavors of Responsibility

by

Susanne Meyer

Department of Philosophy
Duke University

Date: _____

Approved:

Walter Sinnott-Armstrong, Supervisor

Jennifer Hawkins

David Wong

Michael Corrado

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Philosophy in the Graduate School
of Duke University

2016

ABSTRACT

Flavors of Responsibility

by

Susanne Meyer

Department of Philosophy
Duke University

Date: _____

Approved:

Walter Sinnott-Armstrong, Supervisor

Jennifer Hawkins

David Wong

Michael Corrado

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Philosophy in the Graduate School of
Duke University

2016

Copyright by
Susanne Meyer
2016

Abstract

In this thesis, I will assess the question of what it means to be responsible for an action, and attempt to unify the various concepts around the conception of answerability. Focusing on that notion, I argue, will give a better understanding of the other proposed conceptions of responsibility: that of attributability and that of accountability. I will come up with three connected definitions corresponding to these notions, and show that once we refine our concept of responsibility, it is possible for an agent to be responsible in one way, and not another. I demonstrate this contention on the test case of the psychopath.

Keywords: moral responsibility, answerability, psychopathy

Dedication

To my husband, Birkir. In every sense of the word, you made it.

Table of Contents

Introduction	1
Chapter 1: Background.....	2
The Tripartite Theory of Moral Responsibility	7
Conclusion.....	14
Chapter 2: Answerability	16
Working Definitions of Answerability	17
Five Types of Questions — A Quick Overview	20
Imaginability	22
Alternative Options.....	25
Motivation	26
Justifiability	31
Personal Justifiability	33
Blind Spots.....	34
Interpersonal Justifiability.....	37
Blind Spots and Interpersonal Justifiability.....	39
Impersonal Justifiability	41
Conclusion	43
Chapter 3: Attributability.....	44
Examples.....	51
Negligence	55
Unrepentent Sinners	58
Subjects of Attributability.....	60

Addicts	62
Chapter 4: Accountability	65
An Alternative Conception of Accountability	66
Voluntariness	69
Smith and Shoemaker's Legacy	71
Discourteous People.....	76
Accountability Without Attributability.....	78
Attributability Without Accountability.....	79
Interlude.....	81
Chapter 5: An Example: Psychopathy	83
Misconceptions About Psychopathy	87
The PCL-R.....	90
Flavors of Psychopathy	96
Verdicts - Redux	98
Shoemaker and Psychopathic Responsibility.....	103
Psychopathic Responsibility, Susanne Style.....	110
Accountability	114
Conclusion	116
Works Cited	118
Biography	125

Acknowledgements

I am extremely grateful to my advisor, Walter Sinnott-Armstrong, and to Jennie Hawkins, without whose last-minute help and support I would surely have thrown in the towel.

Introduction

This dissertation is about what it means to be morally responsible. This is a heavily contended and difficult issue, and one on which opinions among philosophers (as well as the general public) are divided. In the chapters that follow, I will attempt to systematically assess what moral responsibility means. In the process, I will argue that there are different senses in which one might or might not be responsible for an action. Distinguishing between them will allow us to break the question of moral responsibility into smaller “chunks”, and thus allow us to give a nuanced answer to whether, and in what ways, agents are responsible for their actions.

In order to both motivate and situate my discussion in its larger context, I will begin by providing a bit of background on the larger philosophical debate about what it means to be morally responsible for some thing — in other words, what criteria an attitude or action must meet for an agent to be responsible for it. I will then proceed to discuss in some detail the different senses of moral responsibility that I see as, together, constituting the concept. I will end by attempting to suggest a solution to whether, and in what sense(s), one particularly hotly disputed group of agents are morally responsible for their actions: namely, psychopaths.

Chapter 1: Background

The most prominent criterion for moral responsibility is what philosophers have variously called the volitional or the control views of moral responsibility. Versions of this view are

actually better understood as a cluster of distinct views which share a common assumption, namely, that choice, decision, or susceptibility to voluntary control is a necessary condition of responsibility... what ultimately makes an attitude attributable to a person for purposes of moral assessment is that it is connected in some way to her choices: she made choices in the past which led to the development of the attitude in question, she made a choice in the present to endorse or 'identify with', or she has the ability to modify her attitudes through the choices she makes in the future. It is the connection to past, present, or possible future choices that makes it appropriate to attribute (some) attitudes to persons for purpose of moral assessment, according to this volitionalist accounts. (Smith, 2005, p. 238)

Volitional accounts thus include the hierarchical model described in the work of Harry Frankfurt (1971), according to which we are responsible for an attitude or action if we endorse (or do not endorse) it from the perspective of our second-order desires.

Frankfurt famously writes:

It seems to be peculiarly characteristic of humans that they are able to form what I shall call "second-order desires", or desires of the second order. Besides wanting and choosing and being moved *to do* this or that, men also want to have (or not to have) certain desires and motives. (p. 67)

Frankfurt believes that certain agents (whom he does not want to call persons, but rather calls wantons) fall outside of the boundary of responsible agents because they do not have such second-order desires: "The characteristic of a wanton is that he does not care about his will. His desires move him to do certain things, without its being true of him

that either that he wants to be moved by those desires or that he prefers being moved by different desires.” (p. 11) Frankfurt, then, draws the boundary between responsible agents and those who are not responsible on the basis of a higher-order endorsement of one’s own volitions.

Mark Ravizza and John Martin Fischer offer a different, but still volitional, view. According to them, an agent is responsible for an action or attitude if the agent is reasons-sensitive with regard to it. This means that the agent would change their mind about the action or attitude’s desirability if they were confronted with sufficient evidence to do so. What is required for responsibility is “that there is some possible scenario (or possible world) in which there is sufficient reason to do otherwise, the agent recognizes this reason, and the agent does otherwise.” (1998, p. 44) Excluded from ascription of responsibility, according to Fisher and Ravizza, are therefore those who are immune to reasons to change their behavior, or those who — although they recognize the reasons to change their behavior — are unable to do so. An example of the former might be agents whose behavior is not dictated by reason, but entirely by instinct. By contrast, a representation of the latter is the drug addict who is very much receptive to reasons to change their behavior, but is simply unable to do so. It must be noted that Fischer and Ravizza’s account is much more complex. I am here offering only its very bare bones in order to show that the account is, in fact, a control view of responsibility: agents are in control of their actions insofar as they would change them in response to what would be (to them) good reasons.

Gary Watson (1996) believes that the proponents of volitional theories of moral responsibility miss a significant amount of common moral practice. He advocates for what he calls the 'self-disclosure' view. According to this view, an agent is responsible for an action when their actions express — or disclose — the 'deep' or 'real' self: "Actions which express ourselves in the required sense are free actions, *whatever their farther causes may be.*" (1996, p. 227, emphasis added). Critics of this self-disclosure view (e.g. Susan Wolf, 1990) have argued that the removal of choice or rational control in the production of action undermines the concept of agency, desert, and — most importantly — accountability. Watson readily agrees that 'self-disclosure' views cannot by themselves establish accountability: "While (strict liability aside) attributability in this sense is crucial to the practices of moral accountability, it does not all by itself underwrite it". (1996, p. 229)

However, Gary Watson does *not* believe that 'self-disclosure' falls short of the notion of agency altogether. Self-disclosure, to him, is a bona fide type of responsibility:

I do not agree that the self-disclosure view accounts at best for a 'superficial' notion of responsibility. Moral accountability is only part, and not necessarily the most important part, of our idea of responsibility. The self-disclosure view describes a core notion of responsibility that is central to ethical life and ethical appraisal. In virtue of the capacities identified by the self-disclosure view, conduct can be attributable or imputable to an individual as its agent and is open to appraisal that is therefore appraisal of the individual as an adaptor of ends. Attributability in this sense *is* a kind of responsibility. (1996, p. 229)

It is important to notice that by turning his focus away from choice and accountability, Gary Watson does not contradict volitional accounts of responsibility, but rather enriches the concept with another *kind* of responsibility:

Attributability has an importance to ethical life that is distinct from concerns about accountability. Responsibility is important to issues about what it is to lead a life, indeed about what it is to have a life in the biographical sense, and about the quality and character of that life. These issues reflect one face of responsibility (what I will call its *aretaic* face). Concerns about accountability reflect another. (1996, p. 229)

Attributability thus draws a boundary around the class of agents to whom it applies: the agent must possess a deep self — and must be able to let their action express that deep self. I shall have much more to say about this in the chapter on attributability.

Other philosophers, for example Angela Smith (2005), actively denounce both the volitional and ‘deep self’ views of responsibility, and argue that responsibility is essentially a matter of answerability: that is, of the ability to answer for one’s actions when answers are demanded. Smith illustrates this shift in the concept of responsibility with an example: She invites us to imagine that she forgets the birthday of a close friend. A few days later, she remembers, and is horrified. Clearly, thinks Smith, her forgetting of her friend’s birthday is not an expression of her volition regarding wishing her friend a happy birthday:

“I did not consciously *choose* to forget this special mayor deliberately *decide* to ignore it. I did not *intend* that to hurt my friend’s feelings even *foresee* that my conduct would have this effect. I simply forgot... most of us would probably say that choice or voluntary control is a precondition of legitimate moral assessment. And yet, as the case above was meant to illustrate, we regularly do hold ourselves and others responsible for

things that do not appear to reflect a conscious choice or decision. " (2005, p. 236)

Angela Smith therefore sets her sights on showing "that our commonsense intuitions do not, in fact, favor a volitionalist criterion of responsibility, but a rationalist one." (p. 237)

She writes:

When we praise or criticize someone for an attitude, it seems we are responding to something about the content of that attitude and to facts about its origin in a person's prior voluntary choices, or to facts about its susceptibility to influence through a person's future voluntary choices. More specifically, it seems we are responding to certain judgments of the person which we take to be implicit in that attitude, judgments for which we consider her to be directly morally answerable. (2005, p. 251)

Smith makes herself even clearer: "to say that a person is responsible for an attitude, on the rational relations view, is simply to say that that attitude is, or should be, sensitive to her evaluative judgments and that she therefore can properly be asked to defend or justify it." (2012, p. 267)

David Shoemaker, with Smith's blessing, characterizes her position in the following way: "To criticize an agent morally for something is implicitly to express a demand to her to justify herself..." (2011, p. 605) For Angela Smith, thus ... being responsible for ϕ is most importantly a matter of being answerable to others for ϕ ." (2011, p. 603).

Excluded from Angela Smith's conception of responsibility are two groups of 'agents': firstly, there are the people who act without really knowing, or caring, why they do what they do — Frankfurt would call them wantons. There are such people, and in the chapter on attributability, I will present an example. Secondly, as the name

implies, answerability is all about being able to give *answers* about why one behaved as one did: even if one knows one's motivations, it is still not a given that one can verbalize it, and provide an answer to the question of why one did what one did. Declaring answerability as the centerpiece of responsibility thus presupposes a relatively high threshold of cognitive abilities.

The discussion of the most prominent theories of responsibility thus leaves us with three separate notions of what it means to be responsible: According to the deep-self theorist, being responsible is a matter of standing in a certain relationship of ownership to an attitude: it must be attributable to one. By contrast, according to Smith, being responsible has to do with whether one is answerability for one's action, i.e. whether one is an appropriate target for demands that one justify one's actions. Finally, the volitionalist, believe that one is accountable for an action if performing — or not performing — that action was under one's control.

The Tripartite Theory of Moral Responsibility

David Shoemaker believes that a truly comprehensive theory of responsibility must accommodate *each* of the three notions. (p. 603) This is because it is possible to be responsible in one way, but not in another, for the very same action.

[Smith's] story conflates attributability and answerability, which are actually distinct conceptions of responsibility... [Also], this story conflates two importantly distinct conceptions of responsibility, namely, answerability and accountability. (2011, pp. 603-604)

In order to demonstrate his contention that an action might be attributable to an agent without them being answerable for it, Shoemaker invites us to imagine the mother of a

son who has rightfully been convicted as a serial killer, and is about to be executed. The mother knows — thoroughly believes with her innermost thoughts and feelings — that her son is a terrible person without redeeming features, and that the world is better off with him dead. Nevertheless, on the day of the execution, the mother cannot help feeling dejected. Shoemaker writes: “Emotional commitments are sometimes formed solely in virtue of a ‘mineness’ relation... Where one's attitudes are grounded in those sorts of emotional commitments, they float free of the sort of reasons agents typically take to justify them”. (2011, p. 611) In such cases, argues Shoemaker, it is reasonable to attribute her feelings to the mother, but it would be unreasonable to hold her answerable for them — i.e., to ask her to explain why she has them. She would not be able to so, given her own opinion on the worthiness of her son as a human being¹. What we have, in such cases, is attributability without answerability.

Shoemaker next turns his aim to demonstrating that there are actions for which an agent might be answerable but not accountable. In his 2011 paper, accountability arises when an agent flouts the ‘standards constituting the relationship’. (p. 620) It is not present, by contrast, when the agent merely doesn’t perform an action which his opponent would have liked them to perform, but which would be supererogatory. As Shoemaker puts it,

¹ Shoemaker’s example relies on the (in my opinion) somewhat contentious assumption that ‘mineness’ is not reason enough to justify emotional attachment.

[o]n the one hand, there are standards defined in terms of the claims the parties have on one another, and on the other hand, there are standards defined in terms of the hopes the parties have for one another.... The distinction just offered is essentially the distinction between expectations of good will and expectations of not being treated with ill will. (Shoemaker, pp. 621–622).

He elaborates:

To be answerable...is to be susceptible for assessment of, and response to, the reasons one takes to justify one's actions. To be accountable, on the other hand, is to be susceptible to being held to account if one flouts relationship-defining demands. (2011, p. 623)

Agents thus remain answerable for actions that demonstrate a lack of what Shoemaker calls good will, but fall short of actually demonstrating ill will. By way of example, Shoemaker introduces us to a husband — George — who every year gives his wife — Martha — flowers that she hates (p. 620). She hints at her dislike for the flowers, but her husband never catches the hint, and continues to give her the same type of flowers. Not surprisingly, she is disappointed and perhaps even a bit angry at his insensitivity. It is, according to Shoemaker, perfectly reasonable for the wife to hold the husband answerable — that is, it is reasonable for her to ask him to explain what reasons drive him to buy her flowers every year. It is not, however, reasonable for her to hold him to account, because “[w]hat it must mean to hold someone to account 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.” (p. 623) Giving bad gifts does not warrant such sanctions. Had he cheated on her and thereby clearly violated the rules of their relationship, however, she would have been entitled to hold

him accountable for that behavior — cheating on her would have constituted flaunting the defining characteristics of their relationship.

Critics such as Angela Smith (2011) have pointed out that Shoemaker's example does not demonstrate what he takes it to demonstrate: there is no difference in *kind* between a spouse's obligation to attempt to give one's partner gifts that make them happy, and remaining faithful. Rather, thinks Smith, what Shoemaker describes is simply a difference in the seriousness of the transgression:

In both cases, it seems to me, George is simply answerable for his conduct, because that conduct reflects his evaluative judgments. The differing reactions warranted by his conduct have nothing at all to do with his bearing a different sort of responsibility for what he has done in each case but rather with the responsibility for what he has done in each case but rather with the relative seriousness of the failures his conduct reveals in each case. (p. 587)

The case of Martha and George is not unexceptional in this way. Smith believes that there are in principle no differences in kind among obligations — no dividing line that separates misdeeds from violating the demands for a relationship and mere desiderata in that relationship.

... the sorts of moral responses Shoemaker rightly identifies throughout his paper as key elements of our actual moral practices — aretaic appraisals; feelings of disappointment, resentment, or indignation; modification to our attitudes and actions in response to perceived relational impairments; explicit acts of reproach or censure — are all of a piece, in my view, in the sense that they belong to a single continuum of moral responses that all share the basic precondition of answerability. (p. 589)

I am inclined to agree with this criticism about Shoemaker's chosen example, and indeed his distinction in kind between relationship-defining and relationship-bettering actions,

but do not think that this disagreement threatens the distinction between answerability and accountability: We can still derive this distinction from cases in which it is necessary to perform an action that is, in isolation, morally questionable, but that is on a particular occasion performed in the pursuit of a morally valiant aim. Imagine that you walk across a sweltering parking lot and see me shattering one of your car's windows with a tire iron. You run up to me yelling: "What are you doing?!" This, I contend, is a demand that you justify your actions. When you reach the car, however, you notice that your indoor kitten has somehow managed to sneak into the vehicle, and is close to expiring from the heat of the day. Once you see this, it is unlikely that you will hold me accountable: you will be unlikely to sanction me for smashing your window (you might even reward me). In such cases, I claim, I am answerable without being accountable.

In his 2015 book *Responsibility from the Margins*, Shoemaker refines his distinction between answerability and accountability by stating that answerability is a matter of 'quality of judgment' while accountability reflects 'quality of regard' (p. 11-12). What this means, very roughly, is that answerability "is something agents have in virtue of their of their ability (in principle) to "answer for" their actions, to respond to others' demands for justification by citing their judgment about the worth of some reason over others." (p. 27-28) While "judgments take as their object reasons... regard takes as its object agents" (p. 98). Shoemaker claims that it is easiest to understand what it means to show regard for an agent by considering what it means to *fail* to show regard for an agent. In order to define such a failure, he avails himself of Aristotle's opinion:

“[S]lighting is the actively entertained opinion of something as obviously of no importance...” (p. 93), and adds Hobbes’ statement that “...every man looketh that his companion should value him, at the same rate he sets upon himself, and upon all signs of contempt, or undervaluing, naturally endeavours, as far as he dares... to extort a greater value from his contemners, by damage, and from other, by example.” (p.93)

Quality of regard, then, has to do with the avoidance of slighting others, and valuing them as one values oneself. Shoemaker’s final distinction between answerability and accountability, then, evaluates different things: the former evaluates an agent’s judgment, while the latter evaluates the regard an agent shows their fellow agents. The implication is that it is possible to have excellent abilities to decide between the worthiness of reasons for action without showing proper regard for other agents, and, conversely, show great regard for other agents but be terrible at judging the relative worthiness of actions.

Angela Smith rejects the idea of a tripartite conception of responsibility. In 2012, she writes

[M]y aim in this essay is to argue that our moral practices do not, in fact, embody three different conceptions of moral responsibility and that ‘responsibility as answerability’ is indeed the only kind of moral responsibility there is... What Shoemaker calls ‘responsibility as attributability’, I will argue, is not a form of responsibility at all, and what he calls ‘responsibility as accountability’ is just responsibility as answerability in contexts where an agent has violated specific moral obligation to others. (p. 584)

The reason Angela Smith does not regard attributability as a bona fide type of responsibility is that she believes that there are many psychological features of agents

that form part of their deep selves, but for which we do not want to hold them responsible in *any* way. Smith submits for consideration the OCD sufferer. The OCD sufferer's behavior clearly helps to explain her behavior in the moral as well as in the practical domain: "An agent who suffers from this condition often cannot have normal relationships with others, for she is constantly obsessed with keeping clean or checking over and over again to make sure she has completed certain tasks". (p. 584) Smith agrees that such a person is open to a variety of aretaic judgments, such as that they are neurotic and compulsive. Nevertheless, argues Smith: "[M]ost of us, I submit, do not think of such agents as responsible for this condition or for the behavior it motivates". (p. 485). She elaborates: "And if we go on to ask why we do not think the agents described above are morally responsible for their psychological conditions... I believe the answer is that we do not view them as answerable for these conditions." (p.585). According to Smith, then, our attitudes are only properly attributable to us in a sense that implies responsibility if they are potential targets for demands for answerability: "[when] we are open, in principle, to these sorts of demands with respect to your attitudes, they are attributable to us in the way required for attributions of responsibility in the most basic sense." (2005, p. 271)

In dismissal of accountability as a separate type of responsibility, Smith recalls her belief that transgressions against other agent's expectations are a matter of degree, and not of kind (as illustrated by her response to the Martha/George example). She concedes that the moral responses on which Shoemaker insists as the key to our moral

practices are indeed that, but believes that they occur on a continuum. It thus makes little sense to hold an agent *only* answerable for a lesser offense, but accountable for a more serious one. In either case, thinks Smith,

To say that I am 'morally responsible for ϕ ' is just to say that I am 'answerable' for ϕ and therefore open to legitimate moral criticism if it should turn out that ϕ violates any moral norms or expectations. (2012, p. 578)

Being responsible, then, is just a matter of being answerable for moral transgression. The two, then, are coextensive: being subject to demands for answerability *signifies* responsibility.

Despite her insistence on the co-extension of legitimate ascriptions of accountability and demands for answerability, Smith grants that there "is still plenty of room for meaningful debate over what it takes for an agent to be 'answerable' in the relevant sense for some thing." (2012, p. 589) This debate is precisely what I aim to provide in the next chapter.

Conclusion

In the chapters that follow, I will attempt to shed light on the disagreement between David Shoemaker and Angela Smith. Ultimately, I will agree with Shoemaker that there are three separate notions that must be taken into account in order to give a full account of what it means to be responsible. I will also agree with Smith, however, that answerability is the key concept, and the one from which the others can be derived. The goal of this dissertation is therefore threefold: Firstly, I will attempt to clarify the dispute between the two philosophers by attempting to sort out precisely what they

mean when they invoke the terms 'attributability', 'answerability', and 'accountability'. Secondly, I will offer my *own* definitions of the three concepts. Finally, using those definitions, I will try to demonstrate that it is indeed possible to be responsible in some senses, and not in others. In order to do so, I will utilize a particularly interesting test case — that of the psychopath.

Chapter 2: Answerability

In the discussion concerning moral responsibility, concerns about accountability usually take center stage. From a practical perspective, this certainly makes sense — after all, concerns about moral responsibility usually equate to questions about whether or not we can, reasonably and fairly, hold a person to account for their actions. Similarly, as indicated in the last chapter, the question of what conditions have to be met in order for an action to be attributed to an agent has received much philosophical attention — likely, I suspect, because most people regard attributability as a prerequisite for accountability. This is not to say that the issue of answerability has been neglected: in fact, a few years ago, a discussion between prominent philosophers Angela Smith and David Shoemaker published in *Ethics* (2011-2012) engaged with the topic in great depth. Despite their analyses of the concept of answerability, I will attempt to show that it is still hazy, and that it can be difficult to discern precisely what Smith and Shoemaker are actually saying. In order to clarify this issue, I will offer several interpretations of what ‘answerability’ might mean, and attempt to map Smith’s and Shoemaker’s writings onto these interpretations.

Getting clear on the correct interpretation of the concept of answerability is important. If answerability is to be one of the three types of moral responsibility, we need to know precisely what it means to be answerable for an action — otherwise, it will be impossible to assess whether a particular agent is actually responsible *in that sense* for a particular action. This point becomes even more poignant in the context of Angela

Smith's "aim ... to argue that our moral practices do not, in fact, embody three different conceptions of moral responsibility and that 'responsibility as answerability' is indeed the only kind of moral responsibility there is". (2012, p. 576). Nevertheless, the exercise of refining the definition of answerability is not merely valuable for its own sake: In the chapters to follow this one, I will attempt to demonstrate that if the concept of answerability is framed with appropriate attention to its nuances, the concepts of attributability and accountability — at least equally important in the dialogue about responsibility — can be derived from its various interpretations.

Working Definitions of Answerability

Despite allocating a prominent place in her account on the concept of answerability, it is remarkably difficult to pin down Angela Smith on a precise definition of the concept. In 2005, she introduces her 'rational relations' view, according to which attitudes are the result of corresponding underlying attitudes. Therefore, "since these attitudes are linked with our judgments in various ways, we can, in principle, be asked to defend them and to give up or modify them if an adequate defense cannot be provided" (p. 270). In 2008, she broadens her scope to include actions into the category of events that require a rational connection to underlying judgments. Her stance on what makes agents answerable for their actions, however, remains unchanged: "[t]o say that an agent is morally responsible for some thing... is to say that that thing reflects her rational judgment in a way that makes it appropriate, in principle, to ask her to defend or justify it" (p. 369). By contrast, and by implication, it is *not* appropriate to ask

someone to defend or justify an action if the action does not reflect the agent's 'rational judgment'. Such an agent is not an appropriate subject for the demands for defense or justification.

It is imperative here to note the difference between being an inappropriate target for demands for justification, and *failing* the demands for justification. As I understand it, in the first case, Smith's 'rational implication' between evaluative judgment and action is absent — in other words, the action does not reflect an underlying judgment. A defense or justification would be impossible because there are no underlying judgments to which the agent might refer. By contrast, in the case of failure to defend one's actions, these actions are connected to corresponding underlying judgments and express those judgments, but the action (and by extension, the underlying judgment itself) does not meet with the approval of those who ask the agent to justify themselves. It is in such cases that Smith believes that it is appropriate for agents to "modify" their attitudes and actions (2005, p.270).

David Shoemaker is more thorough in his discussion of the concept of answerability. In his latest book *Responsibility from the Margins* (2015), he provides his own definition of answerability. He does not, unlike Smith, rely on a connection between underlying judgments and actions/attitudes. Instead, his account is more lateral, as it were:

[To be answerable], an agent's attitudes must be generally governable by the agent's judgments about why they are more worthy than (some relevant) others, i.e., judgments about 'instead of' reasons. An agent is answerable for some specific attitude or action just in case the agent could

in principle cite his or her 'instead of' reasons for having the attitude or performing the action... (p. 82)

The concept of "instead of" reasons is eminently important because

... if one is genuinely incapable of seeing the reason in favor of anything other than ϕ , one is obviously incapable of including them in one's deliberations or judging as to their relative worth, and so, it would seem, one could not give a sufficiently robust answer to the answerability demand about one's doing ϕ , even if one takes oneself to have justifying reasons in favor of ϕ . (p. 76)

Note that Shoemaker's explanation of the importance of contrastive reasons begins with a reference to the ability to 'see' reasons. In its own right, the question of what it means to 'see' has received a fair bit of attention, but little consensus (see, for example, Hare, 1952). In addition, in the current context, it is difficult to discern how 'seeing' a reason for something is connected to answerability — after all, *seeing* a reason doesn't necessarily mean that one can *cite* it when one is asked to do so. I assume that this is why, in a very elegant move, Shoemaker treats 'seeing a reason' as a mere condition for the *real* capacity of which one must be capable in order to be answerable: namely, including that reason in one's deliberations. Shoemaker has thus given us a practical interpretation of a somewhat elusive concept, and one we can use in our discussion of answerability. While one might not be able to cite reasons one 'sees', one can presumably cite reasons that are part of one's deliberations.

Smith and Shoemaker disagree, then, on what makes an agent answerable for an action: for Smith, subjects are answerable for an action or expressed attitude if it expresses their underlying judgments. For Shoemaker, an agent is answerable for their

actions if they are able to take into account — and into deliberation — alternative reasons for acting differently from the way in which they ultimately act. In the following few sections, I will discuss the different interpretations of the concept of answerability that give rise to this disagreement.

Five Types of Questions — A Quick Overview

Presumably, in their most general form, demands for an agent to answer for their actions may be represented by the question “Why did you do x ?” I believe that there are several different ways in which to cash out this question: I suggest that there are five interpretive approaches. The distinctions I draw are my own, and — to my knowledge — have never been discussed anywhere or by anyone. In this section, I will very quickly introduce each interpretation, and then discuss them individually in more detail throughout the chapter.

On the first understanding of “Why did you do x ?, what the interlocutor is looking for is a possible causal *explanation* of the action. That is, they want to see whether or not the agent can cite considerations that connect the action to an underlying evaluative attitude that underlies it. The question could be reformulated as “What could *possibly* make you do x ?” I suggest naming this interpretation the question of *imaginability*. A second understanding of answerability is concerned with motivation rather than with imagination. On this interpretation, the question “Why did you do x ?” does not target some theoretical connection between an underlying evaluative judgment

and an action, but rather inquires to know why this particular agent committed *this* particular action at this particular time. This is the question that targets motivation.

The third understanding of “Why did you do *x*?” contains a subtle but very important shift: The interlocutor is no longer interested in what *explains* an action, but rather in what purportedly *justifies* it. On this interpretation, the interlocutor’s goal is to find out what the agents themselves think makes it morally acceptable to perform the action they performed. This type of question targets personal justification — namely, it asks why *that* agents acted in that way under those particular circumstances. Fourthly, when we ask someone why they did what they did, we might be looking for justification in terms that would convince the *interrogator*, instead of the agent, that the action was justified. On this conception of answerability, we are looking not only for a justification of the action that satisfies the agent’s own personal moral beliefs, but rather for a justification that satisfies the interrogator’s. This is the question that aims at uncovering why the agent believes that their action is justified by the standards of a moral authority.

Finally, one might take agents to be answerable to an objective standard of behavior. This is what some moral objectivists believe: to them, *right* or *wrong* exist independently of temporal, geographic, and ultimately fallible human considerations. To a moral objectivist, then, if one deviates from the standard of what is objectively right, one might be called upon to justify doing so not with reference to the *true* moral standard. This type of question, then, aims at a justification in terms of an objective moral standard.

In the following sections, I will discuss each conception of answerability in turn, and relate them to the writings of Smith and Shoemaker. Doing so, I hope, will help to clarify the stances of those writers.

Imaginability

On the first possible interpretation of what we might be asking when we ask an agent to answer for their actions, the answer we seek takes the form of an explanation: we want to know why *anyone* would do the thing they did. We are, in other words, looking for a set of beliefs and desires that *could* make the action intelligible. We are, in this case, not really interested in the specific causal chain that caused the action — we are too perplexed for that. When we ask questions of imaginability, we are simply trying to figure out how an action might make sense; to put it into a possible, but not necessarily actual, cognitive context. Consider, for example, that you are watching TV, and you come across a program that depicts a procession of men and women moving through the streets of a medieval town while violently flagellating themselves (a scene reminiscent of Monty Python's *Life of Brian*). Assuming that you are a psychologically and religiously naive person who knows nothing about Christian (or any other, for that matter) rites of penitence, you might well be horrified and wonder aloud why *anyone* would do something like that. Your friend, who is in the room with you, might then explain to you that self-flagellation is regarded as an important ritual of purification by some religious groups. This might make sense to you, and you might now understand why someone might come to think it a good idea to repeatedly flagellate themselves.

Alternatively, you might simply remain mystified — surely, whatever one’s religious views are, such gratuitous violence cannot please *any* deity. In the former case, your demand to know why *anyone* would act in such a way has been met: a set of beliefs that might explain the action was offered to you. In the latter case, your demand has not been met: a good enough answer was not offered, and you still don’t see why *anyone* would flagellate themselves.

Notice that the actual strength of the explanation matters little here:

Answerability demands are about the interrogator, and thus what determines whether or not such a demand has been satisfied is whether or not the interrogator accepts it as a possible explanation. The explanation that is given does not even have to be the best explanation the agent can give for the behavior: Imagine that you see someone who is in the process of getting a tattoo that depicts Kim Kardashian. You might well wonder why anyone would do such a thing. Your friend, once again, explains to you that getting a tattoo of Kim Kardashian is actually a popular wager in silly bets, and that the person receiving the tattoo must have lost such a bet. It is a plausible explanation, and you might well accept it, but it is, of course, not the *best* explanation: The most likely explanation for the action is that the person getting the tattoo simply likes Kim Kardashian. Nevertheless, the somewhat less likely explanation still satisfies your demand to know why anyone would do as the agent did: the action is coherent given a certain set of beliefs.

A further thing to note is that questions of imaginability might equally apply to actions that the agent did *not* perform. Consider a situation that I have often observed while teaching Intro to Philosophy. I typically teach a version of Philippa Foot's (1967) Trolley problem. The setup I usually use is fairly close to Foot's basic form of the thought experiment: A trolley is loose at the top of a hill and cannot be stopped before it reaches the bottom of the hill. Halfway down the hill, the track diverges into two parallel tracks. On one side, five people are tied to the track. On the other, only one person is tied down. The students are told that they know nothing about any of the people tied to the tracks: nothing about their ages, personal achievements (actual or potential), levels of happiness, etc. — they are simply numbers. The students are then asked to imagine that they are standing at the divergence next to a lever they can pull to change the trajectory of the trolley. Right now, the lever is set so that the trolley barrels down the track to which five people are tied, killing all of them. The question is whether or not it would be morally right (permissible or imperative) to pull the lever. Opinions on what the morally correct action is are usually divided: some students insist that it is wrong to interfere with the trajectory of the trolley, and to actively cause the death of the person who is currently set to survive. Usually, however, the majority of the students believe that it is morally allowed to pull the lever, and save the maximum possible number of lives. Typically, students of either group cannot understand how *anybody* could consider the opposite action morally right: it seems so obviously wrong to minimize the damage, or so obviously wrong to “play God”, respectively. Both sets of

students, then, are unconvinced why anyone would act differently from themselves with regard to the opposite verdict to the problem.

Finally, it should be noted that demanding that someone explain why anyone would do a certain thing is not appropriate to all actions. This is because some actions are by nature done without any reason whatsoever, and there can thus be no imagining why anyone would do them. Think, for example, of Tourette's syndrome. Tourette sufferers certainly *do* things: they have tics that consist in repeatedly making certain noises or physical movements. Nevertheless, it would be unreasonable to ask why *anyone* would act in this way. There is no reason to discover — the tics that characterize Tourette's are simply neurological anomalies.

Alternative Options

The picture gets more complicated when we take into consideration the fact that we usually do one thing *as opposed to* a number of other things we could be doing. We flagellate ourselves instead of going to confession. We get a tattoo of Kim Kardashian as punishment instead of taking the Ice Bucket Challenge. Thus, what the question "Why would anyone do such a thing?" *really* means is "Why would anyone do such a thing *instead* of all these other things they could do?" Take an example from Kent Kiehl's recent book *The Psychopath Whisperer* (2015). Kiehl tells the story of an agent (who happens to be psychopath) who has been invited to a party to which he is to bring his own beer. Once at the convenience store, the agent realizes that he has forgotten his wallet at home. So, he grabs a baseball bat and severely beats the store clerk, steps over

his injured body, and grabs the beer he has come for. When we ask of his actions “Why would anybody do what he did?”, we really mean “Why would anybody beat the store clerk *instead of* going home to get his wallet and pay for the beer?” Juxtaposing an agent’s actions with the actions he *might have* performed instead provides a much more informative picture of the agent’s psychological makeup than looking at the action he actually *did* perform in isolation. In his 2013 paper “Consequentialism, Cognitive Limitations, and Moral Agency”, Dale Dorsey stresses the importance of comparing different potential courses of action in order for putative agents to assert their agency: “Thus to perform an action as an agent it is not sufficient to see some reason or other to perform the action. Rather, one must see what I shall dub a ‘contrastive’ reason. A contrastive reason to ϕ is a reason to ϕ rather than any other action ψ one might perform”. (2013, p. 13) This point is reminiscent of David Shoemaker’s definition of answerability. The ability to imagine alternative courses of action disqualifies one from understanding the reasons others might have for their behavior. This, of course, is the case for young children and the mentally impaired, whose range of potential actions are often severely limited. By contrast, being able to imagine the possible reasons for the courses of action includes an agent among the class of action who can truly examine each other’s behavior.

Motivation

Another sense of asking “Why did you do x ?” aims to uncover the agent’s (purported) motivation, rather than merely some possible explanation for why anyone

might act the way they acted. I say ‘purported’ because motivation is notoriously difficult to discern by conjecture as well as by introspection. We are therefore not looking for *actual* motivation, but only for what the agents themselves take to be their motivation.

Usually, questions that aim to uncover why a specific agent performed a specific act are more precise versions of the question why anyone would commit that action: when we ask someone to explain their actions, it is implied that we don’t want to hear just *any* possible explanation for their behavior, but rather *the* motivation that actually accounts for how they acted. This is especially true in the moral context, in which we are, after all, concerned with specific acts committed by particular agents — we don’t just want to hear an explanation for the sake of an explanation, but rather need to figure out how to react to an agent and their actions in light of their stated motivation. From a practical perspective, then, questions of what an agent can imagine and questions about what motivated them appear closely related. From a conceptual perspective, however, the ability to imagine why anyone would perform an action and the capacity to vocalize why *one* did are quite different in that they require different capacities from agents: Imaginability is about conceivability of reasons. This says very little about the actual psychology of the agent in question. The capacity to verbalize one’s motivations, by contrast, is about introspection, which tells us about the psychological layout of the agent, but little about their capacities for imagination.

It is, of course, possible to be imaginative without being introspective. Take, for example of the first case, the protagonist of Albert Camus' novel *The Stranger*, Mersault. The novel's pivotal scene shows Mersault walking along the beach as he runs into a man, referred to only as 'the Arab', with whom Mersault and his friends have been in conflict in previous scenes. Mersault is on the verge of heatstroke, and dazed by the brightness of the sun. When the Arab flashes a knife at him, Mersault shoots him. Although the Arab dies immediately, Mersault fires four more shots into his body. After Mersault is arrested, he is questioned about why he killed, and in particular overkilled, the Arab. Mersault sincerely considers the question, but cannot come up with a satisfactory answer:

Next, without any apparent logical connection, the magistrate sprang another question.

"Why did you fire five consecutive shots?"

I thought for a bit; then explained that they weren't quite consecutive. I fired one at first, and the other four after a short interval.

"Why did you pause between the first and second shot?"

I seemed to see it hovering again before my eyes, the red glow of the beach, and to feel that fiery breath on my cheeks — and, this time, I had no answer.

During the silence that followed, the magistrate kept fidgeting, running his fingers through his hair, half rising, then sitting down again. Finally, planting his elbows on the desk, he bent toward me with a queer expression.

"But why, why did you go on firing at a prostrate man?"

Again I found nothing to reply.

The magistrate drew his hand across his forehead and repeated in a slightly different tone:

“I ask you ‘Why?’ I insist on your telling me.”

I still kept silent. (pp. 41–42)

Mersault is not holding out on the police — he really does not have an explanation beyond repeated references to the brightness and heat of the sunshine, which are hardly reasons. The action is not unprecipitated — there *has* been conflict between Mersault and the Arab and their respective friends. It is also not particular incongruous against the background of Mersault’s character — Mersault is very detached from the events of his own life and the people in it, regards morality as meaningless, and makes his decisions completely independently of the approval of others. Nevertheless, under the influence of light and heat, he falls into a quasi-trance and becomes a piece of flotsam drifting along on whatever currents run through, or past, him. He does not rationally deliberate on his actions, is not sufficiently invested in them to ensure that his actions are solidly motivated. Rather, he is a passive bystander in his own life, riding waves of nihilism and absurdity. It would be irrational of the police to persist in asking him to explain what motivated his actions — there must be an explanation, a convergence of Mersault’s character and the circumstances of the day, but Mersault has no cognitive access to it (or at least does not seek cognitive access to it). While this means that Mersault is not an appropriate target for demands to reveal his motivations, it does not automatically mean that he is also an inappropriate target for questions about why anyone would act as he did: given the historic unrest between Algeria’s Arab and French inhabitants,

exemplified by the conflict between Mersault's friends and the group of Arabs on the beach, it seems only reasonable to assume that Mersault could easily give a reason for why a French citizen might kill an Arab, even though that is not *his* reason. Excused from answering questions about their motivations, then, are agents who do not have the necessary introspective capacities. It seems, by contrast, difficult to imagine how anyone might be privy to their own motivations, but at the same time lack the imagination why anyone would perform it: if *one* has a reason to do something, than at least *someone* has a reason to do it. One's own description of one's motivation, then, would simultaneously and necessarily serve as an answer to the question why *anyone* would perform that action.

The concept of motivation, I suggest, is part of Angela Smith's description of the eligibility for demands for answerability. Recall that for Angela Smith, to be answerable means that the attitude or action in question possesses a 'rational link' to an underlying evaluative attitude, and it is this underlying attitude that motivates the action of attitude.

The view I am putting forward takes as its starting point the idea that some of our mental states [and actions]² are linked to particular judgments in such a way that, if one sincerely holds a particular evaluative judgment, then the mental state [or action] in question should (or should not) occur. (2005, p. 253)

² In 2012, Smith adds actions to events that are tied in this rational relation to underlying judgments.

In 2012, she adds the short but illuminating assertion that “[The agent] is simply answerable for his conduct, because that conduct reflects his evaluative judgments”. (p.587) It appears, therefore, that it is the underlying evaluative attitude that provides the ‘rational link’ that motivates an agent’s attitudes and actions.

Justifiability

We are now undertaking a significant shift in the concept of what it means to ask an agent why they performed an action: We are no longer merely interested in why anyone would perform a particular action, or why this specific agent did. The question one asks when one demands answerability for an action is not just “Why would anyone do what you did?” or “What motivated you to do what you did?”, respectively, but rather “Why is what you did morally permissible?” We are, in other words, asking a normative rather than merely a descriptive question.

Both Angela Smith and David Shoemaker regard justification as the goal of demands for answerability. In 2012, Shoemaker characterizes what it means to be answerable in the following way: “[t]o be answerable, I believe, is to be susceptible for assessment of, and response to, the reasons one takes to *justify* one’s actions” (623).

In *Responsibility from the Margins*, Shoemaker elaborates:

... suppose I did kick your sister, but my response to your answerability demand is ‘Because I intended to kick her’. This sort of answer may *explain* my action — explain what caused it, what got my limbs in motion — but it is not an answer to the right sort to your demand, which has an implicit reference to *justification*. You want to know what considerations in favor of kicking your sister I thought were worth acting upon... (2015, p. 88)

It is here that the issue of contrastive reasons becomes important: in order to provide justification, it is imperative that an agent have the ability to take into account contrastive reasons, as discussed in the section on imaginability:

... if one is genuinely incapable of seeing the reason in favor of anything other than ϕ , one is obviously incapable of including them in one's deliberations or judging as to their relative worth, and so, it would seem, one could not give a sufficiently robust answer to the answerability demand about one's doing ϕ , even if one takes oneself to have justifying reasons in favor of ϕ . (p. 76)

Indeed, "to be answerable, an agent's attitudes must be generally governable by the agent's judgments about *why they are more worthy than (some relevant) others*, i.e., judgments about 'instead of' reasons." (p. 82, emphasis added)

Angela Smith concurs that the goal of demands for answerability is to uncover justification, not just explanation:

... to say that an agent is morally responsible for some thing is to say that that agent is open, in principle, to demands for *justification* regarding that thing. To blame an agent or criticize an agent morally for something, then, always embodies (at least implicitly) a demand to her to *justify* herself... (2012, p. 578, emphasis added)

Shoemaker and Smith, then, agree that demands for answerability have as their target justifications of an action. Nevertheless, both of them rely on versions of the question "Why did you do x ?" that are aimed at explanation in order to identify who is, and is not eligible, for such demands for answerability. For Shoemaker, as we have seen, an agent can only really be expected to justify their action if they can imagine, and take into their deliberation, an alternate course of action they might have pursued instead. For Angela Smith, the concept of motivation is involved in who is an appropriate target for

demands for answerability: in order to be answerable for a particular action, that action must be motivated by an underlying evaluative judgment, with reference to which the agent justifies the action.

Personal Justifiability

The issue of justification necessarily raises the question of the target of the justification: demands for justification might consist in demands that the agent defend their action against the background of some external standard of conduct. Alternatively, agents might be asked to justify themselves against the moral code of the person who subjects them to demands for justification. Or, they might simply constitute demands that agents tell us why they themselves believe that their actions are justified. In the remainder of this section, I will discuss the third type of justification, which I call *personal justifiability*.

In the case of *personal justifiability*, what we ask is that agents justify their actions against the background — as Smith would put it — of their own evaluative judgments. Thus, when we ask an agent to personally justify an action, we aren't just asking "What motivated you to do what you did?", but rather "What makes *you* think that it was morally permissible for you to do *x*?" It is critical to stress here that it is the agent's own moral code that provides the point of reference against which the action is to be justified, *not* that of the interrogator. We might ask such questions simply out of anthropological curiosity — this occurs in contexts in which we have no interest or reason to judge the moral codes of others. Consider, for example, the Western tourist who travels to Japan in

order to explore a new culture. Much of Japanese society, to this day, is governed by the bushido code: the traditional samurai code of conduct that is based on honor and discipline. Western tourists might marvel at the resulting behavior of Japanese people, and might even ask them (or perhaps a tour guide) to explain why they do the things they do. Most tourists will, however, understand that Japan's unique history shaped the local code of conduct, and — unless it greatly offends them — refrain from judging it against their own values. (Certainly, a few will be unable to criticize the local moral code from the presumed superiority of their own: these tourists slip into the sense of answerability I will discuss next.)

It should be noted that some philosophers believe that personal justifiability is the only form of moral assessment we are ever justified in conducting. An extreme example of this position is, famously, Jean-Paul Sartre, who believed that there was no right or wrong beyond what the individual agent whole-heartedly embraced. It is easy to see that if this is the case, any moral inquiry can only be, as Sartre calls it, moral anthropology.

Blind Spots

Not every agent is an appropriate target for demands for personal justifiability — and more importantly, not every agent is an appropriate target for demands for personal justifiability *for every action*. Consider, for example, Shoemaker's example of a population of aliens who believe that it is morally wrong to walk on grass:

In particular, they think it immoral to walk on the grass, precisely because of what it does to the grass: it bends and breaks it. It is

intrinsically bad, they claim, for this sort of organism to be bent or broken, and they purport to ground this claim on their understanding of what it is like to be a blade of broken or bent grass. (2012, p. 625)

A visiting earthling, Shoemaker claims, would be unable to “empathize” with grass and “enter into its being-space”: he would, in fact, have no idea what the aliens are talking about (2012, p. 626). Were the earthling to tramp across the grass, it would presumably be unreasonable for the aliens to demand that he justify himself: where the moral status of grass is concerned, the earthling suffers from a ‘moral blind spot’. Indeed, asserts Shoemaker: “One could, therefore, have various local blindnesses, that is, be able to recognize the “instead of” reasons only in some normative domains but not others. This would render one an answerable agent only within those former domains.” (2015, p. 82)

Not everyone is convinced that Shoemaker does, in fact, have a blind spot with regard to the moral status of grass. Matthew Talbert (2012), for example, argues that Shoemaker knows very well what it means to show empathy — he just does not believe that grass deserves this type of consideration. As Talbert puts it:

[T]he problem with Shoemaker’s example is that that it obscures the difference between two distinct ways in which agents may have difficulty abiding by a moral norm. An agent who is aware that her action violates (what others take to be) a moral norm might violate the norm because he has difficulty seeing the force of the moral considerations that are offered in support of it. Another agent, who feels the force of the consideration offered in support of the norm, might still fail to abide by it because she reasonably (but wrongly) rejects the claim that her action violates the norm. (2012, p. 564)

What I take this paragraph to mean is that it is, indeed, possible to have moral blind spots, and not to see the moral considerations in favor of a certain action. What is also

possible, however, is that one is well acquainted with the moral consideration in question *per se* (in this case, respect for other beings), but does not believe that one needs to follow them in a particular case. Talbert believes that in the case of the aliens and their empathy for grass, Shoemaker takes himself to have a blind spot about the moral treatment of grass. Talbert, by contrast, argues that Shoemaker has no such blind spot — he knows what empathy is — but believes that the aliens falsely apply it to grass. One of Shoemaker’s quotations actually supports Talbert’s point: “what the aliens deem immoral about grass-trampling I merely see as stupid: I am just incapable of viewing the grass’s bending and breaking *as giving me reason of any kind*. So as I chortle about the aliens’ ridiculous moral beliefs, I tramp across the grass.” (2011, p. 626, emphasis added). Notice that Shoemaker no longer claims that he does not understand what the aliens mean by “empathizing” with grass — he now appears to simply find it ridiculous that they apply this consideration to it, and finds that he has no reason to follow their beliefs. Shoemaker, then, does not *really* have a moral blind spot when it comes to treating other beings with respect — he just thinks that the aliens are wrong in their application of moral consideration to grass.

There are, however, agents with true, and considerable, blind spots. Think, for example, of children: Young children begin their lives as creatures of instinct. The development of action-guiding codes of conduct, let alone moral codes of conduct, takes years and much practice (see, e.g., Kohlberg (1981); Piaget (1932)). This implies that the personal justifiability of children is not an all-or-nothing matter: rather, when we want

to know whether or not a child is properly subject to such demands for a particular action, we need to pay attention to the status of their (moral) development: as they age and learn, children move from a state of global exemption in terms of personal justifiability to a state of local excuses with regard to certain issues, to — ideally — agents with not restrictions on their demands for personal justifiability. In the interest of foreshadowing, my last chapter will discuss another group of agents who have a sizable moral blind spot: namely, psychopaths.

In order to be an appropriate target for demands for personal justifiability, then, an agent must have a sufficiently robust moral code to guide their behavior, and they must be able to verbalize it. In the case of moral blind spots, agents do not have such moral codes, and are therefore excused from having to justify their actions that fall within that moral blind spot.

Interpersonal Justifiability

It might appear that where Shoemaker is concerned, personal justifiability fits his concept of answerability: one is answerable for one's actions insofar as one can justify them to oneself as more worthy than their alternatives. However, in revealing passages in *Responsibility from the Margins*, Shoemaker asks us to imagine a race of superhuman or aliens who purportedly recognize all sorts of "reasonish" facts that we cannot — in his example, extra reasons not to drink Pabst Blue Ribbon beer. When they ask us why we drink it anyway, claims Shoemaker, "we are not answerable *to them* for drinking it if we

are unable to recognize the relevant (to them) 'instead of' reasons for not drinking it"

(p. 77). He adds:

Answerability is, at the end of the day, about agents' liability for evaluation of a certain sort. We do not make our answerability demands just to hear others' 'instead of' reasons full stop, as in, 'Oh, I see, so you killed him because you thought it would be more fun than helping him? Great, thanks for your quick reply!'. Instead, hearing the reasons sets the stage for a response... (p. 77).

What these passages imply is that what Shoemaker has in mind when talking about answerability is not merely the ability to justify one's actions against *one's own* moral code — not merely moral anthropology as discussed in the previous section. Rather, the passages suggest that answerability, at least sometimes, involves justifying oneself against the moral codes of others.

Angela Smith concurs with the contention that answerability is a matter of answering *to others*. This is, it appears, because she regards answerability as synonymous with accountability (2012), and accountability is by its very nature an interactive matter. Recall her statement that "[t]o say that I am 'morally responsible for ϕ ' is just to say that I am 'answerable' for ϕ and therefore open to legitimate moral criticism if it should turn out that ϕ violates any *moral norms or expectations*." (2012, p. 578, emphasis added) The 'moral norms or expectations' she has in mind are, of course, those of other agents.

What is needed, then, is a further sense of answerability, which I shall call *interpersonal justifiability*. According to this sense of answerability, when we ask an agent to answer for their actions, we do not just require that they justify them with reference to

their own moral code. Rather, we want them to justify their actions with regard to *ours*. It should be noted that interpersonal justifiability is an inherently relativist concept: it makes no claim about the source of the interrogator's³ moral code, but fixes it as the required point of reference regardless of that source. What is required for demands for interpersonal justifiability, however, is that the interrogator takes themselves to hold a position of moral superiority over the agent who is held interpersonally answerable. It should come as no surprise, then, that demands for interpersonal justifiability are primarily used in situations in which we are interested in judging, and perhaps subsequently punishing, other agents.

Blind Spots and Interpersonal Justifiability

From the foregoing paragraph, it seems that the conditions for interpersonal justifiability are akin to those of personal answerability: in order to be an appropriate subject for demands for answerability of either type, the attitude or action to be explain must fall outside of a moral blind spot. What I also believe, however, that it is possible to have a moral stance — i.e. no personal blind spot when it comes to justifying one's position in one's own terms, and yet have a blind spot when it comes to justifying that same moral stance in the terms of another's moral code. Consider an example. In a 2005 edition of *The Daily Show*, Jon Stewart discussed the moral status of gay marriage with

³ 'Interrogator' might here stand for an individual or a group of some kind, as long as they represent a unified moral code.

Republican presidential candidate Rick Santorum. In summary of their discussion, Stewart remarked:

It is so funny; you know what's so interesting about this is ultimately you end up getting to this point, this crazy stopping point where literally we can't get any further. I don't think you're a bad dude, I don't think I'm a bad dude, but I literally can't convince you. (Sinnott-Armstrong, 2010, p. 331)

Although both men seem to be cognitively capable to engage in discussion, and both seem to have very strong personal moral code to which they repeatedly referred, it seemed to Stewart that neither man would ever be able to justify his position in the terms of the other's moral code. Each one of them found the other's moral standards mystifying *as* moral standards — Stewart did not get the idea of the sanctity and purity of heterosexual marriage, and Santorum did not understand how something so clearly (to him) disgusting could be part of anyone's code of morally acceptable practices. An inability to explain one's moral position on a certain issue, then, can appear even in cases in which one can easily justify that moral position in terms of one's own moral code.

What this means, I think, is that in general, the same capacities are necessary in order to justify a moral position against one's own moral code, and against someone else's moral code. As we have seen, what is necessary, firstly, is that such a moral code exist, and that the agent have the cognitive capacity to verbalize the relevant beliefs. However, I also believe that if the agent's moral code and that of the interlocutor are formulated in terms that are so alien to one other that each interlocutor cannot even see

them as reasons for action in any instance, it is impossible for an agent to answer demands for interpersonal justifiability. This example differs from Shoemaker's aliens example in that, in Shoemaker's example, the earthling was familiar with the concept of empathy, and accepted it in general as a reason for action — just not where grass is concerned. In my example, by contrast, Steward and Santorum regard believe that the other's moral code is *never* justified, and so cannot use it to justify their actions with regard to *any* being.

Impersonal Justifiability

There is one final type of answerability that I would like to mention for completeness' sake. This is what I call *impersonal justifiability*. The concept of impersonal justifiability interprets the question "Why did you do *x*?" as "What objective truth justifies your action?" On this interpretation, then, an agent would be answerable for an action only if they are able to justify it against an impersonal moral standard: a moral standard that exists apart from local and temporal human endorsement, i.e. a morally *objective* standard. I contend that the conditions of impersonal justifiability are much the same as those of interpersonal justifiability: They require that an agent justify their actions against a moral code that transcends their own. In the case of interpersonal justifiability, that moral code is that of the interrogator. In the case of impersonal justifiability, that moral code is an objective truth (or a perceived objective truth) of which the interrogator takes themselves to be a representative. A good example of a demand for impersonal justifiability is the sacrament of confession as (historically)

practiced in the Catholic church. During confession, the sinner is asked to reveal their deeds that violate the dictates of God's will. Having confessed their transgressions, the agent is met with corresponding sanctions, usually in the order to pray a certain number and type of prayers. This corresponds to the criticism or reactive responses an agent might receive in the case of interpersonal demands for justifiability.

Similarly, I think that the limitations on an agent's ability to answer demands for impersonal justifiability are akin to those on their capacity to answer demands for interpersonal justifiability: if the impersonal requirement is of a *kind* that does not figure in the agent's personal moral code, it will be difficult for them to fulfill those demands (as demonstrated by the familiar vs. formal address example). If, by contrast, the moral requirement is of a type that also figures in the agent's personal code of conduct, but is required by the purported objective truth to a greater degree or as applied to a wider range of agents, the agent is, according to the representatives of the objective truth, a fitting target for demands for impersonal justifiability. That is why Catholics pay penance: the goal is that in the future, agents will be humbler than they already are, more charitable than they already are, more forgiving than they already are, etc.

There is no indication in either Smith's or Shoemaker's writings that their concept of answerability sets the bar as high as impersonal justifiability. In fact, Angela Smith ties demands for answerability to cases in which 'moral norms and expectations' (2012, p. 578) are violated. This choice of words, though not decisively so, appears to refer more appropriately to human-made standards than to objective truths. Shoemaker,

by contrast, offers up various examples in which agents are (or are not) answerable to other types of agents (e.g. aliens upon whose grass earthlings tread, superhumans who know of additional reasons for refusing to drink Pabst Blue Ribbon), but never mentions answerability with reference to an objective moral code. Therefore, I contend that while impersonal justifiability *is* a possible interpretation of the original question “Why did you do x ?” it is not what either Smith or Shoemaker has in mind. Furthermore, it is conceptually similar enough to interpersonal justifiability to be regarded as a version thereof, and thus not require its own in-depth discussion.

Conclusion

Having read about the distinction between the different possible interpretations of the question “Why did you do x ?”, the reader might nevertheless remain anxious to know which of the five ways in which to cash out the original question constitute demands for answerability. I hope to have demonstrated that the answer to this question is: the ones that involve justification rather than only explanation. Of these demands for answerability, it is interpersonal answerability that corresponds to the Smith and Shoemaker’s concept of answerability. Nevertheless, these concepts of answerability depend on an agent’s capability to answer questions of imaginability and motivation.

Chapter 3: Attributability

Thus far, we have been focusing on the concept of answerability and its various interpretations. For some philosophers, this focus is justified, and indeed exhausts the discussion of responsibility. Remember, for example, Angela Smith's assertions that "'responsibility as answerability' is really the only kind of moral responsibility there is," (2012, p. 576), and that "[w]hat Shoemaker calls 'responsibility as attributability', I will argue, is not a form of responsibility at all." (2012, p. 584)

David Shoemaker disagrees with this claim. In "Attributability, Answerability, and Accountability" (2012) he attempts to establish that there are, in addition to answerability, two further types of responsibility that can, and ought to be, conceptually separated from answerability: attributability and accountability. In order to gain a full conception of responsibility, one must include all three. Attributability, then, is a type of responsibility in its own right.

Gary Watson (1996) concurs with Shoemaker on the importance of attributability as a form of responsibility. He, too, believes that attributability is a real type of responsibility. He stresses the importance of aretaic appraisal, i.e. the practice of judging an agent's character as falling short of some moral standard without holding them accountable for this failure. Thus, I might find it regrettable that a colleague does not behave more virtuously without blaming her for doing so: had I been raised in the way she was raised, had I gone through what she went through, I would probably act as she

does. Such aretaic appraisal, Watson argues, nevertheless constitutes a real sense of moral responsibility:

... aretaic appraisal involves an attribution of responsibility. To adopt an end, to commit oneself to a conception of value in this way, is a way of taking responsibility. To stand for something is to take a stand, to be ready to stand up for, to defend, to affirm, to answer for. Hence one notion of responsibility—responsibility as attributability—belongs to the very notion of practical identity. (Watson, 1996, p. 271)

What is important to note is that the dispute is not whether or not agents make these types of judgments — all the philosophers mentioned in this chapter agree that they do. The dispute between Smith on the one hand, and Shoemaker and Watson on the other, is only whether those judgments constitute a type of responsibility. I would argue that for our purposes the issue matters very little, and is largely semantic. What *is* the case is that attributability, as Shoemaker and Watson define it, is a separate concept from answerability and accountability. Shoemaker and Watson's assurances that attributability *is* a separate type of responsibility, and Smith's that it is *not*, both make the concept worthy of investigation.

Attributability, according to David Shoemaker, is based on what Susan Wolf (1990) calls the 'deep self.' Although the precise nature of this deep self has been much disputed, Shoemaker identifies two general approaches: Some philosophers, notably Harry Frankfurt, believe that the deep self expresses itself in the active endorsement of one's own volitions from a high-order perspective. The problem with such an account, as Gary Watson (1996) famously points out, is the threat of an infinite regress: the high-order volitions are themselves just volitions, and it is thus hard to see why they should

be any more representative of one's deep self than one's first-order volitions. In response to this difficulty, Watson proposes a conception of the deep self that locates the deep self in an agent's evaluative stances, rather than in their volition.

We might say that an agent's values consist in those principles and ends which he—in a cool and non-self-deceptive moment—articulates as definitive of the good, fulfilling, and defensible life... we all have more or less long-term aims and normative principles that we are willing to defend. It is such things as these that are to be identified with our values. (1996, p. 25)

Shoemaker believes that both Frankfurt and Watson's conceptions of the deep self are overly intellectual (2015). If one limits one's conception of the deep self to cognitive judgments of what desires are worth having, or what aims are worth pursuing, one risks missing out on entire classes of 'cares' that are not quite so intellectual, but rather visceral. Shoemaker thus proposes a version of the deep self that encompasses both intellectual and emotional commitments. He calls this view

Ecumenical Attributability (EA):

An attitude is attributable to an agent just in case it expresses the agent's deep self, that is, it is causally dependent on, and its content is harmonious with, at least one of the agent's cares, or care-commitment clusters. (2015, p. 75)

Shoemaker defends his Ecumenical deep self theory by claiming that the 'cares' or 'attitudes' in question in his view have two distinct sources — and therefore so does the deep self. (2015, p. 50) He cites Gary Watson as the proponent of one of these sources: there are "those principles and ends which [the agent]—in a cool and non-self-deceptive moment — articulates as definitive of the good, fulfilling, and defensible life."

(Watson 2004, p. 25). On the other hand, Shoemaker points out, “many representative attitudes [of the deep self] have their ultimate source in our cares, where these [...] are dispositions to respond emotionally in sync with the fortunes of the cared-for object.” (Shoemaker, p. 51) Thus, “[c]ares and commitments both seem to express character. They are both forms of *mattering*.” (2015, p. 51) What Shoemaker takes himself to do, then, is correct a perceived weakness in both Watson’s and Frankfurt’s account: he adjusts their accounts so as to save them from being overly intellectualized.

Presumably, some of these cares — namely the emotional ones — are merely attributable to the agent: they form part of their deep self, but are beyond demands for answerability. Recall, by way of Shoemaker’s own example, the mother of the about-to-be executed serial killer who, despite knowing that he is the scum of the earth, cannot help feeling bereft on the day of the execution. By contrast, some of the commitments Shoemaker attributes to the deep self are more intellectual. It would make sense to ask the agent to justify these types of commitments. The Ecumenical View thus encompasses both cares that are merely attributable to the agent, and cares for which they are answerable — attributability and answerability, on Shoemaker’s concept of the deep self, sometimes overlap. Moreover, an agent might be *accountable* for parts of their deep self — recall from chapter one that accountability, on volitional accounts, depends on making a choice about endorsing, and potentially indulging, a particular attitude or character trait (or, for that matter and in the interest of Shoemaker’s Ecumenical View, a more emotional care). *Prima facie*, it seems reasonable to believe that one might be

accountable for part of one's deep self: an agent might spend a considerable amount of effort on cultivating a certain aspect of their deep self: think, for example, of the gang member who purposely engages in violence to 'toughen himself up', and embed a penchant for violence deeply into his own deep self. It is therefore not clear which one of the criteria for responsibility is at work in identifying the deep self: the commitments that are attributable to the agent by virtue of his intellectual endorsement, the cares that stem from his emotional cares, or the character traits he has cultivated — or all three.

I propose a simpler way to track the deep self. According to my suggestion, in order for an action, attitude, or character trait to belong to an agent's deep self, and thereby make it attributable to that agent, the agent does not simply have to know what *motivated* it, but must also think that it is *justified*. Note that there are a number of different things that might be attributed to an agent: feelings, actions, character traits, attitudes, etc. There are several caveats to add to this definition from the get-go: Most importantly, regarding oneself as justified for committing an action comes in degrees. Watson, for example, regards actions as attributable to an agent if they deem them 'good' — and a single sentence later if they deem them 'defensible' (Watson, 2004, p. 25). There is, of course, a humongous difference between the two: In the former case, one takes a positive stance regarding the worth of a purported value; in the later, one takes a negative stance: the action is simply not morally forbidden. I shall follow Gary Watson in this: being attributability responsible for something is a dimensional concept. In other

words, one is attributability responsible to the extent to which one regards an action or character trait as justified.

Moreover, I believe that it is common not to be able to articulate one's attitude toward a certain thing. This is a point that distinguishes my account from that of Watson and Frankfurt: those accounts depend on the agent's taking a specific and explicit stand on potential cares and commitments. Mine does not. An example will illustrate the point. Imagine that I am in Toronto, ON, after dark, and I take out my trash. As happens far too often, there is a raccoon digging through my bin, and spreading the contents all over the yard. Suppose that I am usually an animal lover, so it would not be fair to attribute an unmitigated attitude of hatred toward animals to my deep self. I even like raccoons sometimes, as long as they stay away from my trash bin — they can be pretty cute. However, having gone through this same hassle more times than I can count, I throw a kick toward the raccoon (knowing, of course, that it is far too fast to be caught by it). What should we say about my deep self in such a situation? I have already indicated that I believe that it would be unfair to simply attribute hatred of animals (or raccoons, specifically) to me. I like them a little bit, and I hate them a little bit. I cannot articulate how much I like or despise them — and yet both attitudes have a certain stability to them. It is not that my attitude to raccoons swings back and forth like a pendulum from adoration to hatred — if that were the case, we could hardly say that either attitude stems from my deep self, since my deep self would then be ever

changing. Rather, both my stances toward raccoons stem from my deep self — my deep self-attitude toward them is mixed⁴.

Secondly, it must be remembered that answerability, as discussed in the previous chapter, is about the questions an agent can answer — or could answer, if they were posed to them. The criterion for attributability is much weaker: in order for an action to be attributable to an agent, they need not be able to vocalize their commitment to certain values — attributability may be judged in terms of non-verbal behavior. Agents do not take themselves to be justified in a certain action if they, for example, attempt to hide it, attempt to change it, or express remorse about it (provided always that they are sincere about their efforts and are not simply trying to avoid the law). This point ties into the one made in the last paragraph: one might wonder *how much* struggling is necessary for an agent to escape attribution of a certain character trait. The answer is that there is no answer: as I have attempted to argue, the attribution of character traits is dimensional. The more evidence (external or private) there is that an agent finds themselves unjustified in their actions, the less justified we are in attributing it to the agent's deep self.

⁴ This is not to say that it is impossible to have a single and completely homogenous attitude toward a thing — but that is rare.

Examples

My proposal is best understood by way of examples. Take, firstly, the overwhelmed mother— let's call her Amanda. Amanda's children have spent the past few hours bickering and goading each other into fights. Amanda has told them numerous times, as calmly and gently as possible, to settle down, to engage in more sedate activities, to go to separate rooms until they are ready to behave, etc. After the tenth unsuccessful plea, Amanda can no longer keep her irritation in check, and yells at the kids to stop their shenanigans. She might even include some very mild epithets, but epithets nonetheless. Immediately, however, she feels bad — yelling at one's kids is not only bad parenting, but, to her mind, it is also morally wrong. Amanda therefore sees herself as having violated her own moral code: she can *explain* how she came to yell at her children, but she cannot *justify* it. Because of this, I contend, the action does not reflect her deep self, and is not attributable to her.

There are two points to stress at this point. Firstly, the fact that the action is not attributable to Amanda does not automatically mean that she is not accountable for it. The criterion for accountability is different from the criterion for attributability, as I will argue in the next section. Secondly, remember that attributability is about violating one's personal moral codes. It is well possible that witnesses to the incident, for example Amanda's husband — let us call him Bob — thinks that she is well justified in yelling at the children after having tried, and failed, to get their attention and cooperation in gentler ways. But this does not make a difference to whether or not the action is

attributable to Amanda — what matters is her *own* perception of its justifiability. She rejects its defensibility, so it is not attributable to her. Notice, however, that in case Bob's reasoning that while staying calm is generally one of the main roles of a parent, *sometimes*, when the children are exceptionally bad and the parent is exceptionally tired or stressed, there may be *some* justification in yelling at them — sometimes one has to do what one has to do in order to be heard. If Amanda accepts this reasoning, she might well come to see herself as *a little bit* justified in losing her temper — and if that is so, we may attribute the action to her to a corresponding degree.

Contrast Amanda's situation with a situation that has recently appeared in the local news. On September 20, 2015, Durham resident Alan Tysheen Eugene Lassiter threw his three children into a pond. The oldest child, at seven years old, was able to escape, but his three- and five-year-old sisters remained in the pond until an off-duty police officer pulled them out of the water. Sadly, three-year-old Calista Lassiter died on September 23, 2015 (the five-year-old survived, and was released from the hospital in good condition on September 24, 2015). Shortly after throwing the children in the water, Lassiter called 911, and told the dispatcher that he attempted to drown his children "because Child Protective Services 'wanted to take them away from me'". CPS was thus trying to take "the rest of the, the little happiness I had" ("3-year-old dies after father allegedly threw her into Durham pond: WRAL.com," n.d.). It is obvious that the father suffers from a number of mental pathologies: in addition to the pedophilia, there is — even to the non-psychiatrist — a strong suspicion of a personality disorder. Most parents

would not only realize that in the circumstances described in the story, the children would be better off living apart from their father, but would also prefer their children's survival to their deaths even in the event of their being taken away by CPS, no matter for what reason. Nevertheless, it seems that Lassiter is not only able to cite his motivation for throwing his children into the lake, but also regards himself as fully justified in his actions. In his phone call to 911, he says: "[I] blame CPS for this. They are trying to take people's kids over dumb [expletive], and this is what you get!" ("3-year-old dies after father allegedly threw her into Durham pond: WRAL.com," n.d.). What else, to Lassiter's mind, could, or should, a father do under such circumstances?

There is a further facet to the Lassiter case, and one which may *not* be attributable to him by my own definition: It turns out that CPS was involved in the family's life because the father, in his own words, "was dealing with some pedophilia things". Actions precipitated by those impulses resulted in the investigation of Lassiter's fitness to parent young children. However, while Lassiter attempts to justify drowning his children in a pond so CPS wouldn't take them from him, he does *not* attempt to justify his pedophilia. By contrast, he states that he had been attempting to "get some help" for his sexual desires. What this suggests, however vaguely, is that Lassiter realizes that pedophilic sexual desires are not the kind of thing that can be morally justified. They simply *are*, they control and explain his behavior, but that does not mean that he believes them to be justifiable (provided he is sincere in his attempt to seek help). Therefore, I believe that Lassiter is attributability responsible for trying to drown his

own children, but he might not be attributable for his pedophilic impulses. This, I admit, might be a tough pill to swallow: how can a pathology like pedophilia, which plays such a serious role in determining Lassiter's behavior, not be part of his deep self? The first thing I would like to point out in order to stave off this objection is, once again, that the fact that I do not see Lassiter's pedophilia as attributable to him does not, by any means, mean that he is not accountability for it. The second thing is that it is difficult to tell, from the news story, just how much Lassiter struggled against his impulses. If he is in any way disingenuous, it is well possible that the impulses are a part of his deep self.

Consider a clearer, though fictional, example based on the 2004 movie *The Woodsman*. In the movie, a convicted (and unnamed) child molester returns to his hometown after 12 years in prison, moves into an apartment across from an elementary school, and struggles against his reawakening desire to molest children. Although he comes close to succumbing, he ultimately gathers fortitude against his own impulses when his intended victim tells him that her own father abuses her. On his way home from this encounter, the reformed child molester comes across another man, whom he has observed abducting children. In a fit of rage against the other child molester — and likely, against their shared pathology — the main character beats his nemesis badly enough to fracture his jaw. Although both men are pedophiles, or at least have pedophilic desires, I believe that the movie's main character is not attributable responsible for these pedophilic impulses — they are not part of his deep self because he does not believe that he can justify them. Indeed, much of the movie is dedicated to

his struggle against his urges. What *is*, therefore, attributable to him, are his efforts to subdue the impulses that beleaguer him (as well as a willingness to engage in violence).

While we never learn much about the mental life of the Woodsman's nemesis, the blatancy and enthusiasm with which this character indulges his desires suggests that this pedophile does not, in fact, find fault with his own behavior: he appears to believe that he is morally justified in this behavior by at least his own moral standards. If this is indeed the case, then he, as opposed to the main character, is attributability responsible for his desires — they form part of his deep self.

Negligence

It has been suggested to me that my account cannot deal with cases of negligence. If I act negligently, the thought goes, how could this negligence not be attributable to me? The short answer is that to the extent to which it the negligent action expresses a pattern, it negligence is part of the agent's deep self. To the extent to which it is simply a mistake, it is not.

Consider the case of the law-abiding citizen who accidentally runs through a Stop sign (let's call him Charlie) — distracted, perhaps, by his thoughts about a dissertation he must read that night, or simply by tiredness, or by the fact that the Stop sign is very hard to see in the midst of the foliage that surrounds it. Imagine that Charlie is stopped by a police officer and confronted with the offense. His first reaction is the thought (or possibly exclamation) : "Oh my God, I'm sorry." If Charlie is given a ticket, he does not feel unjustly treated — he ran a Stop sign, running Stop signs is against the

law, the law is to be obeyed, and he must therefore pay a fine. Charlie, however, did not run the Stop sign out of a sense of being justified in doing so: As stipulated in the example, Charlie's first reaction is one of contrition — this implies that he does *not* take himself to be justified in running through the Stop sign. If he did, he would — maximally — attempt to argue with the police officer about how Stop signs should be clearly visible, or about how there was no traffic to sharpen his attention, so his running the Stop sign didn't matter since it didn't endanger anyone. Minimally, he would seethe the against the injustice of having been issued a ticket for such a justifiable transgression against the law and wait until the last moment to pay it. As stated before, in order to determine attributability, we do not require that the agent be able to vocalize why, and how, his punishment violates the moral codes by which he abides — it suffices that he chafes against it with a certain level of intensity. Therefore, I contend that Charlie's negligence is *not* a (significant) part of his deep self. This is true even if questions of justifiability never entered his mind when he saw the flashing lights behind him — he never even conceived of the idea that he might justified in his actions. He was simply sorry. Notice, however, that the fact that Charlie is not attributability responsible for running a Stop sign does not mean that he is not *accountable* for doing so. The criterion for accountability is entirely different, and will be discussed in the next chapter.

Imagine, however, that Charlie runs through a Stop sign not only today, but also next week, and the week after that, and the week thereafter. On my account, the one-time offender is not attributability responsible for his isolated act of negligence because

it does not really express carelessness as a significant tendency of his deep self — he does not think his lapse in attention is justifiable. But what should we make of the agent who habitually runs Stop signs — let's call her Dana? It seems to me that there are two possibilities. Either Dana keeps running through Stop signs because she *actually* takes herself to be justified in doing so under certain conditions, no matter how contrite she might seem when he is caught. Alternatively, Dana will take steps to correct her behavior, however minimal — she will promise herself (and the officer who stops her) that she will be more careful in the future. In the former case, I suggest, we may regard Dana's negligence as attributable to her as part of his deep self: she doesn't *really* think she's doing anything wrong. By contrast, in the latter case, I think we cannot attribute negligence to her deep self — rather, her repeat offenses are an expression of forgetfulness and distraction (which we can attribute to her deep self), but not of a belief that negligence in traffic is justified.

There is one more point about my account of attributability that this example underscores: attributability is not about the seriousness of a habit or character trait, but rather about the agent's relation to it with regard to its justifiability: a few pages ago, I declared The Woodsman, a pedophile, *not* attributability responsible for his impulses. In the immediately preceding paragraph, I find a mildly negligent driver attributability responsible for her actions. Attributability really is about the stance an agent takes towards a character trait, *whatever* that trait may be.

Unrepentant Sinners

There is another class of agents about whom my proposal appears to yield counter-intuitive results. The agents I have in mind are those who are fully aware that their passions cannot be morally justified, not even to themselves, but who nevertheless indulge in them with abandon and very little guilt. My account would imply that the actions of such agents are *not* attributable to them, and yet the willingness — and perhaps even glee — with which they engage in them appears to make it difficult to swallow the contention that the actions are not, in fact, expressions of the agent's deep self. Assume, for a moment (and this is historically highly questionable at best) that Vlad the Impaler knew very well that there is *no* possible justification for having hundreds of peasants impaled on stakes in front of one's castle in order to watch them die. Imagine that even Vlad himself thought that this was an abhorrent thing to do — even to *his* mind, there were less brutal ways to maintain his position of power. Nevertheless, suppose that Vlad the Impaler really enjoyed watching his victim's pain; and that, in fact, the wrongness of inflicting such pain added to his pleasure. On the present proposal, it seems that we would have to conclude that Vlad the Impaler is *not* attributable responsible for his passion for impaling whoever he can because he does not think that it is justified. That, however, seems a ludicrous claim to make — how could this passion *not* be part of Vlad's deep self?

There are two potential ways in which I believe one might approach this worry: Firstly, one might simply bite the bullet and accept the conclusion that (our likely

psychologically inaccurate version of) Vlad the Impaler's passion for impaling people is *not* attributable to him. However, it might be argued, what *is* attributable to him in the sense that it forms part of his deep self is an astonishing level of callousness and brutality: our Vlad, in order to act on a fetish he himself believes to be morally wrong, must believe that he is justified in operating beyond the dictates of morality. He might never have thought about it, but if he did, he must believe that he is justified in giving his own desires preference over his moral judgments. And for *that*, he is attributability responsible. In a sense, the particulars of his preferences are arbitrary: instead of watching people being impaled, he might have loved to keep them chained up in his dungeon, or he might have liked to rape and choke them, or he might have been partial to any other one of countless morally unjustifiable actions. What is *not* incidental, by contrast, is the fact that his personal value systems places his passions above his moral convictions. *That* value structure forms part of his deep self.

The other way to address the worry starts out from the opposite assertion: one might be tempted to say that the willing moral offender is, in fact, attributability responsible for his actions. Thus, in our example, Vlad *is* attributability responsible for his passion for watching people die on stakes. This answer to the apparent tension follows an Aristotelian thread, and simply denies that Vlad does, indeed, believe that his actions are morally indefensible: if he did, the thought goes, he would take at least *some* steps to stop himself from acting in ways that he believes are wrong, and he would feel at least *some* guilt at failing to do so. The fact that he does not feel any guilt or attempts

at moderation shows, conversely, that he takes himself to be perfectly justified in acting as he does, no matter what explicit statements he might make. On this solution, Vlad the Impaler is (in terms of his attributability responsibility) the equivalent of the habitual stop-sign runner — who *would* have taken steps to stop the Stop sign running if she really felt it was unjustified. Both suggestions appear to rescue my proposal from the intuitive threat of the willing wrongdoer. I am partial to the latter because it corresponds more closely to my analysis of attributability, but see no prima facie objection to the former line of reasoning.

Subjects of Attributability

What my examples suggest is that whether or not an agent is attributability responsible for a given pathology depends not on the nature of that pathology, but rather on the agent's attitude toward it: whether or not they believe that the actions associated with the pathology are justifiable determines whether or not they are part of the deep self. This also means, of course, that the same agent can be attributability responsible for a mental disorder at one point in time, and not attributability responsible for the same disorder at another time. Consider, for example, the case of Evelyn, an anorexic. In the early stages of the disorder, Evelyn regarded her disorder as an intruder into her otherwise non-disordered mental life. She lamented her own abhorrence of fattening foods: for example, she repeatedly commented on her own inability to eat a Big Mac. She knew she would love it, but could not put aside her worries about the calorie count long enough to enjoy the burger. However, as the disorder progressed, Evelyn

began to rationalize her anorexic impulses: in the later stages of the disorder, she claimed that she could not eat apples that had rested in a fruit bowl next to a mound of shelled peanuts because peanuts are very high in fat, and so the apples must be “infected” with calories by the proximity (personal communication, circa 2001). What I suggest is that Evelyn’s anorexia became attributable to her, as a feature of her deep self, at the same rate as she began to justify, rather than merely to explain, the impulses it caused. Agents can grow into and out of attributability for traits of characters, then.

This is even the case for characteristics an individual has had their entire life. Think, for example, of the Woodsman example discussed a few sections ago. The Woodsman has, most likely, had pedophilic impulses his entire life. Perhaps, at some point, he saw nothing wrong with them, and they were therefore attributable to him. Nevertheless, at some point in his adult life, the Woodsman decided that pedophilic impulses — and in particular the actions they entailed — were no longer a justified feature of his life, and he set out to eradicate them. At *that* moment, they began to lose their status as part of his deep self. Conversely, imagine a kid from a decent home who begins to run with the wrong crowd. This child has never shown any sign of callousness, but after witnessing his mates torture smaller, weaker creatures for a certain number of times, becomes inured to doing so, and begins to think that torturing small creatures is a morally defensible thing to do. As this process unfolds, callousness towards animals becomes, increasingly, part of the kid’s deep self.

Similarly, my account of attributability leaves it irrelevant *how* an agent acquires the character traits that are attributable to them (although, once again, this may make a difference in terms of accountability). Think, for example, of Susan Wolf's second-generation dictator JoJo, who learns his despotic ways from his father's brutal example. JoJo learns well, and his father's lessons come to form part of his deep self — they are attributable to him. Contrast this with the case of, say, Frank, who learns at his grandmother's gentle knee that certain groups of humans deserve certain types of treatments. If Frank grows up to embrace the judgments his grandmother advocated as indeed justified, that belief is *also* attributable to him. Attributability, then, has nothing to do with length for which a belief has been held, or its origins — all that matters is the relationship between an agent and the justifiability of an action or attitude.

Addicts

The issue of addiction is a good illustration of this last point. Chandra Sripada recently wrote a very insightful paper (accepted but not yet published) on the issue of the responsibility of addicts. In it, he assesses the difference between Frankfurt's Unwilling and Willing Addict, respectively, and finds that Frankfurt is correct in finding the first *not* responsible for his addiction, while the latter *is* responsible for his addiction. Sripada begins by referring to Frankfurt's definitions of the unwilling and the willing addict, respectively. The unwilling addict:

hates his addiction and always struggles desperately, although to no avail, against its thrust. He tries everything that he thinks might enable him to overcome his desires for the drug. But these desires are too powerful for him to withstand, and invariably, in the end, they conquer

him. He is an unwilling addict, helplessly violated by his own desires.
(Frankfurt 1971, p.12)

By contrast,

... [The willing addict] is altogether delighted with his condition. He is a willing addict, who would not have things any other way. If the grip of his addiction should somehow weaken, he would do whatever he could to reinstate it; if his desire for the drug should begin to fade, he would take steps to renew its intensity. (Frankfurt 1971, p.19)

According to Sripada,

It is noteworthy that there is little debate about the Frankfurtian position's first claim that the Unwilling Addict is not morally responsible for his drug-directed actions. This is a point of consensus among theorists as diverse in their views about moral responsibility as Frankfurt, Fischer, Wolf, T. M. Scanlon, and R. J. Wallace (p. 5)

Sripada's claim corresponds to the current literature (personal communication, Dr. Rogers, MD, 2016), which suggests that there is such a thing as an addictive personality, and persons who are born with it will sooner or later (usually sooner) succumb to addiction (naturally, they also need the opportunity to become addicts). The disagreement, then, is about whether Frankfurt's Willing Addict is responsible for his addiction. Sripada explores eight arguments in favor of the willing addict's responsibility in order to ultimately demonstrate that Frankfurt's analysis is right.

Sripada acknowledges that neither addict could stop using their drug of choice — or only with great effort — if they chose to do so. Stopping, in other words, would be just as difficult for the willing addict as it would be for the unwilling addict. Therefore, Sripada reports that Frankfurt recognizes that “[c]ontrol based views [...] say that, with regard to their drug-directed actions, the addicts are the same with respect to moral

responsibility – neither addict is responsible” (2015, p. 2). What this means is that if control over one’s action is the criterion by which one assess an agent’s responsibility, we cannot draw a difference in responsibility between the willing and the unwilling addict. However, there another way to draw a responsibility-related distinction between the willing and the unwilling addict: “in Frankfurt’s cases, both addicts lack control over what they do, but they differ in terms of self-expression; the Unwilling Addict does not express his self in action while the Willing Addict does.” (2015, p.2) What this means, then, is that the Willing addict’s addiction is attributable to them because they regard it as a good thing, as a justifiable thing, as a thing that they treasure, and that therefore is firmly entrenched in their deep self. The Unwilling addict, by contrast, resembles *The Woodsman*: his life is dominated by the desire, and the resulting attempts to rid himself of the drug addiction that ruins his life.

These examples, I hope, have served to suggest an alternative way in which to delineate the deep self: whether or not an action or attitude expresses an agent’s deep self depends on the extent to which they embrace it as justified.

Chapter 4: Accountability

The concept of accountability is often regarded as synonymous with that of responsibility — this is, I suspect, because in the moral domain the ultimate question is often whether, and to what extent, to hold agents account for their actions. It is not surprising, then, that the philosophers whose work I have discussed have opinions on what makes an agent accountable for an action. Angela Smith, for example, sees moral accountability simply as (interpersonal) answerability in cases in which the agent has done something wrong: “[W]hat Shoemaker calls ‘responsibility as accountability’ is just responsibility as answerability in contexts where an agent has violated specific moral obligation to others.” (2012, p.584) This seems to me to be an anemic description of accountability.

David Shoemaker is more thorough in his approach to accountability. To him, “to be accountable for something ... is to be liable to being appropriately held to account for it, which is to be susceptible to a range of confrontational responsibility answers” (2015, p. 116) Notice here the difference between *being* accountable, and *being held* accountable: Being accountable is about liability to a range of ‘confrontational responses’ — it does not automatically mean that one is actually being met with such confrontation responses. Being held responsible, by contrast, entails that one is *actually* the target of such responses. The ‘responsibility answers’ Shoemaker has in mind are reminiscent of Strawson’s reactive attitudes. Among them, Shoemaker regards the reactive attitude of *anger* on the part of the victim or their representative as the quintessential indicator of

the agent's accountability (2015): getting angry at someone (or having cause to get angry at someone) is the purest form of holding them accountable. Shoemaker defines anger in the following way: "The syndrome essentially consists in (a) a feeling of heat or aggression (Eatough and Smith 2006), (b) thoughts about slights, and (c) a tendency to prevent or retribution (communicated as such)." (2015, p. 90)

Anger is warranted when another agent has failed to show proper 'quality of regard' to the agent in question. It is difficult to pinpoint exactly what Shoemaker means by 'quality of regard', and so he lets Aristotle have the honor:

So what does it mean to regard other agents? I think we can best understand it by focusing on its poor instantiation — disregard — and looking for the fittingness conditions of its target sentiment, which is what evaluates regard as poor. To do so, we can return to Aristotle's definition of disregard: [S]lighting is the actively entertained opinion of something as obviously of no importance..." (as quoted in Shoemaker, 2015, p. 93)

It needs to be stressed that Shoemaker is not claiming that *other* confrontational attitude such as blame, resentment, and indignation cannot be appropriate responses to such perceived slights — he is merely claiming that anger is the confrontational response that best describes our (actual or potential) reaction to perceived slights to 'quality of regard'.

An Alternative Conception of Accountability

In the following sections, I will attempt to offer a conception of accountability that is, hopefully, somewhat richer than Smith's, and somewhat simpler than Shoemaker's. I will formulate my concept of accountability in the terms I've been using throughout the previous chapters. I am doing this in the hope that doing so will clarify

how accountability relates to the other types of responsibility I have discussed. Since accountability is by nature about how people react to one another, it should come as no surprise that the basis of my definition is the concept of *interpersonal answerability*. Instead of being *answerable* to a moral authority, however, this chapter will assess the conditions for being *accountable* to a moral authority. What I like to offer as a definition of accountability is this:

An agent is accountable to a moral authority for an action when they (voluntarily) violate that authority's moral code of conduct, *despite* their ability to regard that code of conduct as *some* reason for action.

This means, in plain English, that I am accountable to you when I *do wrong* by your moral norms *even though* I am cognitively and psychologically equipped to take into deliberation your reasons for action.

The question what normative principles underlie the moral authority to which an agent is accountable is difficult, and far too complex to address at length in the current context (for a good discussion of the problem, see, for example, Koorsgard, 1996). Nevertheless, I would like to say a few sentences about it. It seems to me that one source of the moral authority of some agents or agencies over others comes, quite simply, from their position in society. This applies, for example, to parents (and children, who have a right to a certain type of treatment), teachers, Home Owner's Associations, and the like. A second source of moral authority comes from an agent's identification with a certain group — in that sense, an agent chooses a moral authority by which to abide: if I join Scientology, for example, I am thereby subjecting myself to the moral authority of that

group. The third way to create a locus of moral authority, as I see it, is simply to enter into a social contract with other agents: whether an agent does so implicitly or explicitly, the rules that foster cooperation among agents take on moral authority for those who live together in a society. Thomas Hobbes puts the point succinctly:

For every man looketh that his companion should value him, at the same rate he sets upon himself, and upon all signs contempt, or undervaluing, naturally endeavors, as far as he dares... to extort a greater value from his condemners, by damage, and from others, by example. (Hobbes, 1998)

What these different sources of moral authority imply is that it is possible to be accountable to one agent or group but not to another at the same time: take the example of Travis Alexander, a devout Mormon — and equally devout party boy (Alexander is famous because he was eventually gruesomely murdered by his girlfriend Jodi Arias, who will feature in the next chapter). Suppose that Alexander goes out on a Saturday night with his buddies — as he often did — stays out until 2 am, and gets inebriated. When he gets to church on Sunday morning, there is much tut tut-ing about his obvious hangover. He is accountable to his congregation; he has chosen to be part of it, and he has violated its moral code despite seeing some reason not to do so (namely, that he swore he wouldn't). At the same time, Alexander is *not* accountable to his friends because he has done nothing wrong by their standards: their entire friendship is built around a mutual love of partying. It is also possible — and probably the most common of cases — that agents are accountable to each other. This occurs when neither has, or claims, moral authority over the other: both, as Hobbes describes, simply want to pursue their endeavors without being hindered by the other agent. Moral authority in such

cases, I contend, is not tied to a particular agent or group, but to the concept of mutual laissez-faire. If *you* violate the moral code of cooperation despite taking into account the fact that it will derail my projects, then you are accountable to me for violating our presumption of cooperation. If *I* do the same thing, I am accountable to you. An example will serve the point: Often, there is strife between neighbors, especially if they live very close together. Some neighbors mean-spiritedly do whatever they can to make their neighbors' existence uncomfortable: they install incredibly bright spotlights that are pointed at the neighbors' bedroom windows, they light fires in their yards knowing that the smoke will drift to the neighboring property, etc. In such cases, although neither agent is in principle morally superior to the other, the mean-spirited neighbor is accountable to his neighbor for violating the implicit social contract that exists between them.

Voluntariness

A further thing to notice is that my definition of accountability contains the condition that an agent must violate their moral authority's edicts voluntarily. This condition is meant to exculpate agents who are forced or compelled⁵ to act against the moral authority that holds sway over them (and which they see as giving them reason to act). Suppose that you are on your way to your car after a long day at work, and out of

⁵ I am using these terms in the usual way: "forced" implies coercion by an outside force, whereas "compelled" refers to internal psychological forces. Doing something voluntarily, then, means that an agent is under the influence of neither.

the bushes jump two thugs with guns. They force you, at gun point, to get into your car, drive to the closest bank, and rob it. You are, presumably, violating your own moral code, for the validity of which you see ample reason — people cannot just go around robbing banks. Clearly, though, we do not want to hold you accountable for your action: you did not *voluntarily* violate your moral code, you were *forced* to do so. The involuntariness of your action, then, undermines your accountability for the action.

The case of an agent who violates the terms of a moral authority out of compulsion is more difficult to imagine, and far more controversial. Consider, for example, the woman to whom nothing is more important than being a good mother to her children, and giving them what is their right — a good home, a mother who is invested in their projects, etc. Suppose, also, that this mother has the great misfortune of being a hoarder: she is compelled to spend all her days hunting for treasures at thrift shops, and fills her house to the brim with junk. She doesn't want to do this, but she simply cannot stand the thought of a great bargain going to someone else. The children feel neglected, and eventually the house is enough of a junk heap that they cannot live there anymore, and must leave their mother to her obsession while they go to live with their father, who has long since left to make his own life. The mother cries herself to sleep every night, knowing that she is betraying everything she believes in, but is up early the next day to hit the flea markets. Do we want to call her accountable? I think not: certainly, she is violating the edicts of her most important moral authority — that of motherhood — but she does not do so voluntarily. Her hoarding is a disease, an

addiction, a compulsion. It seems to me that we ought to treat her with compassion and offer her help, instead of holding her accountable for her behavior (this is why I *hate* the show *Hoarders*). The compulsory nature of her actions undermine her accountability.

Smith and Shoemaker's Legacy

It should be evident that my definition of accountability borrows heavily from both that of Angela Smith, and that of David Shoemaker. Very generally, from Smith, I am borrowing the rough structure of the definition: The concept of being answerable for an action indeed forms the basis for the notion of being accountable for it. Nevertheless, I believe that my concept refines hers: it tells us just which kind of answerability is a precursor for accountability — interpersonal answerability. Furthermore, it is not enough, on my account, to 'do something wrong' in addition to being answerable: one has to do something wrong *in spite of* being able to take into deliberation something *right* in the eyes of the moral authority. This 'in spite of-ness', or insistence on contrastive reason, is a concept I borrow from Shoemaker: on my account, as on his, one can only be accountable if one can take into deliberation reasons for acting differently from how one eventually acts. I cannot be accountable for an action, for example, if it violates a moral norm, but I cannot see *any* reason whatsoever to follow that moral norm. Imagine that I am a teenager, and I am hell-bent on studying drama. My parents try desperately to steer me toward a subject that they see as having a more secure future: economics, for instance, or computer science. But I see *no* reason whatsoever to study either of these

two: the jobs studying economics or computer science will garner me do not interest me in the slightest, and I cannot imagine leading that sort of life. To me, being interested more in art than money, their argument has no force whatsoever, and I am unable to take it into my deliberations when plotting my course of action. But the ability to appreciate instead-of reasons forms a large part of Shoemaker's account: if I cannot imagine why *anyone* would do something (such as study economics or computer science), I do not have the ability to take those purported reasons into deliberation when charting my life, and hence I am not accountable for the decisions I make.

There is a further point of convergence between my definition of accountability and Shoemaker's account of responsibility: note that my definition refers to the *ability* to see other agents' reasons. Shoemaker puts the point in the following way:

Now of course incapacitation is quite different from capacitated failure. Many respond single-mindedly only to the reasons in favor of ϕ , without ever considering reasons for not- ϕ , but that cannot mean they are not answerable when they do so. (2015, p. 77)

This point is important because sometimes we act as the result of some conscious amount of deliberation — like, for example, when I decide to buy a new necklace instead of donating the price thereof to my friend's charity. In such cases, I consciously take the reasons for donating the money to charity into deliberation: it would help someone, it would make me a better person, etc. On other occasions, I act without really thinking about it at all: I might walk without a second thought past a person who desperately needs my help. Were you to stop me and explain that person's need, I would, of course, see the reasons for helping, and would certainly take them into deliberation. As it is, I'm

just oblivious. Nevertheless, I think that in either case I am equally accountable for my actions. This, incidentally, is why we are justified in holding accountable the stereotypical BMW driver who fails to use his signal lights when driving: This is not only discourteous, but actually dangerous: a recent study indicates that not using turn signals actually causes many accidents (<http://philadelphia.cbslocal.com/2012/05/02/new-study-says-failure-to-use-turn-signals-is-a-leading-cause-of-car-accidents/>, accessed October 8, 2016). The BMW driver is accountable to us — he is ignoring the rules of the road that were set to protect everybody, and thus constitute a moral authority. He is, presumably, perfectly able to appreciate the reasons for using turning signals — he just thinks that as the king of the road, he is above the requirement to do so, and thereby violates the rules of courtesy.

Despite these points of agreement, my account of accountability differs from Shoemaker's on a number of important points. Most strikingly, my account makes no mention of 'quality of regard', which is central to Shoemaker's concept of accountability — to him, as I mentioned at the very beginning of the chapter, being accountable to someone equates to the obligation to show 'quality of regard' to them. Shoemaker admits the difficulty of defining this 'quality of regard', and avails himself to Aristotle's definition of slighting someone: "[S]lighting is the actively entertained opinion of something as obviously of no importance..." (cited in Shoemaker 2015, p. 123) In showing regard,

...while we do not expect strangers to care about many of the particular things we do (as we expect friends and loved ones), we typically do

expect a minimal level of shared cares with strangers, cares for the basics we all tend to have as humans, for example, care about avoiding pain, meeting basic nutritional needs, and so on. (2015, p. 103)

While my account regards social contracts of the type Aristotle describes as a source of moral authority in their own right, it avoids the potentially confusing concept of 'quality of regard', which Shoemaker himself is only able to define in the negative. As Angela Smith's own definition of accountability suggests, the simplest and 'cleanest' way in which one might define accountability is through the concept of moral authorities — and that can be done without ephemeral terms like quality of regard.

Another thing to distinguish my account from that of David Shoemaker is that the way I have presented accountability here is completely independent from the concept of accompanying sanctions: as I have presented it, being accountable is simply a matter of *doing wrong by someone* (i.e., acting against the dictates of a moral authority), but doing wrong is not integrally tied to the being the subject of 'confrontational reactions', let alone to the actual reception of such confrontational regards. Although my account does not invoke such a tight connection between being accountability and liability to sanctions, it is of course true that even on my account, violating the dictates of a moral authority is often followed by punishment for doing so. Such punishment often reflects the moral authority's perceived seriousness of the transgression, as well as the level of contrition shown by the transgressor. This is part of the relevance of our discussion of motivation in the chapter on answerability: The motives that offenders cite when asked about the origins of their actions are often used to determine punishment.

Consider two examples: In 1995, Timothy McVeigh — with the assistance of Terry Nichols — bombed the Alfred P. Murrah Federal Building in Oklahoma City. 168 people died in the attack, many of them children who were attending the in-house daycare facility. McVeigh revealed that he deeply hated the national government for its role in the Waco, TX siege in 1993 as well as its handling of the Ruby Ridge incident of 1992. In 2001, McVeigh sent a letter to Fox News in which he made the following statement:

I decided to send a message to a government that was becoming increasingly hostile, by bombing a government building and the government employees within that building who represent that government. Bombing the Murrah Federal Building was morally and strategically equivalent to the U.S. hitting a government building in Serbia, Iraq, or other nations... Based on observations of the policies of my own government, I viewed this action as an acceptable option. (n.p.)

By contrast, in 2012, University of Alberta medical student Christopher Westbrook Mancuso missed 12 shifts at a Winnipeg hospital's emergency room. In an effort to graduate despite not fulfilling his program's requirements, Mancuso forged the signatures of his supervisors to indicate that he had, indeed, attended the shifts he in fact skipped. When he was found out, Mancuso "expressed remorse and acknowledged it was unprofessional to skip the shifts". He admitted that he had acted "in a moment of panic". ("Medical student expelled after caught forging MDs' signatures - Manitoba - CBC News," n.d.)

Both men were punished for their transgressions: McVeigh, of course, was tried, condemned to death, and executed. The harshness of this punishment reflects both the

seriousness of his crime against humanity, and his complete lack of contrition (McVeigh referred to the dead children as “collateral damage”).

(<http://articles.latimes.com/2001/mar/29/news/mn-44250>, accessed Oct, 8, 2016)

Christopher Mancuso, by contrast, faced the Ethics board of his university, and was expelled from his program. The relatively mild nature of his offense and his expression of remorse induced the Ethics board to withhold further sanctions, such as prosecution. These examples show, I believe, that although, such as anger, are not an integral part of my concept of accountability. Nevertheless, we sometimes want to apply sanctions to violations of the norms of social authorities. When we consider what sanctions or punishments to apply, we often refer to what the agent answer when asked what motivated them to do as they did⁶.

Discourteous People

What should we make of people who do not accept that any code of conduct except their own? Alternatively, what do we make of agents who accept the moral authority of other agents or agencies, but find their edicts to be ridiculous, and therefore without any reason to be followed? Think, for example, of the high-level executive who realizes that there is such a thing as a code of courtesy by which agents conduct their lives, but believes that this code does not apply to them? Think, also, of the above-cited

⁶ Other factors also come into play, such as the fact that McVeigh killed many people, and Mancuso hurt only himself.

BMW driver who (stereotypically) neglects to use his turn signals because, in his own mind, he is the king of the road, and need not pay heed to common rules of the road. My analysis might appear that such agents are not captured within the boundaries of accountability: they either do not accept any authority, or do not regard the edicts of that authority as reason for action. But this is not the final story: whether or not these discourteous people recognize it, they are a part of what Hobbes defines in this way: "For every man looketh that his companion should value him, at the same rate he sets upon himself." (Shoemaker, 75-76). In other words, being a member of a corporate structure, or of traffic, entails that one is subject to its moral demands. As Shoemaker reminds us:

Now of course incapacitation is quite different from capacitated failure. Many respond single-mindedly only to the reasons in favor of ϕ , without ever considering reasons for not- ϕ , but that cannot mean they are not answerable when they do so.

Most discourteous people, then, are not unable to see the value of courtesy per se, nor the concept of moral authority that supersedes even their own perceived authority. These agents, wrongly, do not think that they themselves are subject to other moral authority's dictates — courtesy is for important, but only for *lesser* people.

There are, however, agents who are genuinely incapable of accepting moral authorities other than themselves, and reasons that other purported moral codes attempt to impose on them. We shall be talking about them in the next chapter.

Accountability Without Attributability

The fact that my concepts of attributability and accountability are derived from different notions of answerability heavily implies that they are, in fact, independent of one another. I have already hinted at the idea that accountability can exist without attributability — in fact, that idea was the point of the Stop sign example presented in the previous chapter. In that example, we saw that we would not attribute the character trait of negligence to a driver who very occasionally and accidentally fails to stop for a Stop sign because his action is not an expression of taking himself to be justified in running Stop signs — he does *not* take himself to be justified in doing this. Nevertheless, we might hold him accountable for his action: he has broken the moral (and legal) codes of an authority despite being able to appreciate the authority's reasons for making its rules — it is dangerous to run Stop signs.

A similar case may be made for actions committed on a whim: When I was a child, my father, who is a very moral and law-abiding man, once stole something. The item was a \$2 sticker of the Swiss flag which we did not need, and did not even particularly want. He simply walked by it, it caught his eye, and in a moment of exuberance, he pocketed it. My siblings and I, also moral and law-abiding people, were thrilled at his daring — and in roughly equal measures appalled at the immorality of the action. I can guarantee that my father is not the type of person to believe that stealing is morally justified, even if it is just a \$2 sticker — it is still stealing, which is something we shouldn't do. Thus, I do not hold him attributability responsible for the fact: it does not

express his deep self. Nevertheless, my siblings and I clearly held him accountable for the action. He had defied a moral authority (Thou shalt not steal) despite his ability to take into deliberation the reasons for this moral edict.

Akrasia has also been suggested as an example of accountability without attributability – occasions, in other words, on which I simply cannot bring myself to do what I know to be right (e.g. Arpaly and Schroeder, 1999; R. Jay Wallace, 1999). I am skeptical of this particular example: it would seem to me that cases of *akrasia* reflect a conflict among two facets of the deep self, each one of them stable enough to create a significant conflict between the two. But if this is so, and if the agent ultimately chooses the action inspired by laziness or avoidance, it would seem that they are not only accountable for it, but that this action is also attributable to their deep self as an expression of its deeply conflicted nature. *Akrasia*, then, is not a good example of accountability without attributability.

Attributability Without Accountability

Attributability, of course, can also occur without accountability. This occurs in cases in which an agent has a deep self that is strongly dominated by a particular character trait, but is unable to take into deliberation the edicts of a moral authority. For an example of a case that shows, purportedly, that attributability is *not* sufficient for accountability, we need look no further than Susan Wolf's case of JoJo (1990). (Susan Wolf, it will be remembered, is responding to the 'deep self' view, advocated by many philosophers, including Harry Frankfurt (1969; 1971) and Gary Watson. (1996)

JoJo is the favorite son of Jo the First, an evil and sadistic dictator of a small, undeveloped country. Because of his father's special feelings for the boy, JoJo is given a special education and is allowed to accompany his father and observe his daily routine. In light of this treatment, it is not surprising that little JoJo takes his father as a role model and develops values very much like Dad's. As an adult, he does many of the same sorts of things his father did, including sending people to prison or to death or to torture chambers on the basis of whim. He is not coerced to do these things, he acts according to his own desires. Moreover, these are desires he wholly wants to have. When he steps back and asks, "Do I really want to be this sort of person?" his answer is resoundingly "Yes," for this way of life expresses a crazy sort of power that forms part of his deepest ideal. (p. 379)

Nevertheless, Wolf maintains that it is "dubious at best that he should be regarded as responsible for what he does" (p. 379). She thinks this because it "is unclear whether anyone with a childhood such as his could have developed into anything but the twisted and perverse sort of person that he has become" (p. 380). While JoJo's actions really express his deep self, that deep self, because of the non-rational way in which it is was formed, is normatively *insane*, and therefore excludes him from accountability (Wolf, 1990).

It is interesting that not everyone is convinced by Wolf's example: David Faraci and David Shoemaker (2014) introduced the case to their students and found that a majority of them did, in fact, find JoJo accountable for his actions. Faraci and Shoemaker theorize that this is the case because their students believe that JoJo had plenty of opportunity to observe his father's behavior, and to realize that there were different ways in which to treat people. In other words: JoJo adopted his despotic ways *despite* being able to take into consideration a different way of ruling. This, according to my

definition (and the students' intuition), makes him accountable. The issue we are once again rubbing up against (as in the case of Shoemaker's earthling tramping on grass) is the difference between inability to take other points of view into deliberation, and the simple refusal to do so.

While the case of JoJo is therefore a questionable example of an agent whose actions are attributable to them, but who are not accountable for them, we will see a clearer example of such a case in the next chapter.

Interlude

In this chapter, I have attempted to propose a conception of accountability that refines that of Angela Smith, and simplifies that of David Shoemaker. In doing so, I have attempted to rely on the terminology of the previous chapter in order to demonstrate how accountability relates to answerability and attributability. My definition of accountability requires that the agent take into deliberation an authority's moral code when deliberating on their courses of actions, rather than invoking it in justifying their behavior. I hope that the connection between the two concepts is evident.

This chapter concludes the philosophical investigation of the three concepts that, together, make up that of responsibility. So far, I hope to have established two things: I hope to have shed some light on the debate between Angela Smith and David Shoemaker, published in *Ethics* in 2011 and 2012, by attempting to clarify just what these philosophers *mean* when they use terms related to the concept of responsibility. The second thing I hope to have achieved is to show that by refining the central concept of

answerability, we can illuminate the relationships among the different types of answerability I have identified, and ultimately their relationship to answerability and accountability, respectively.

The next, and last, chapter will be less theoretical: I will attempt to apply the different conceptions of responsibility to a test case — that of the psychopath. Doing so, I hope, will reveal that it is possible to be responsible in some ways, and not in others. This is exactly what happens with the psychopath.

Chapter 5: An Example: Psychopathy

In May 2013, Jodi Arias was convicted in Phoenix, Arizona of the first-degree murder of her ex-boyfriend Travis Alexander. During her trial, it emerged that after the killing, Arias sent sympathy flowers to Travis's grandmother, who had helped to raise him. Prosecutor Juan Martinez asked her why she did this, knowing that she herself had killed Alexander. Arias replied: "In retrospect, it probably wasn't a good idea. But I felt that it would be more insensitive to not do anything at all" (Arizona vs Arias, Feb. 25, 2013).

Paul Bernardo is a notorious Canadian serial killer and rapist. He was found guilty of a variety of crimes — including two first-degree murders and two aggravated sexual assaults — in September 1995 and is currently serving a sentence of life without the possibility of parole for at least 25 years at the Milhaven Institution in Bath, Ontario. Bernardo does not accept the contention that he is a psychopath: "I wake up every day knowing I'm not psychopathic. I care about people. I cried during 9/11. I cried during Columbine" (quoted in Makin, 2008, para. 2). The violence involved in the crimes that landed him in prison might appear to contradict this statement, but Bernardo doesn't think so: "If you look back at some of the first [rapes] — in '88, the girl got her arm broken. A lot afterwards didn't. ... You have ups and downs..." (quoted in Makin, 2008, para. 20).

In this last chapter, I will discuss a class of agents whose moral responsibility is hotly disputed: that of psychopaths. I will start with a few examples of philosophers' positions on the responsibility of psychopaths. There are, of course, many more opinions

in the literature than I have the space to discuss, and I inevitably neglect some positions that deserve to be discussed.

Some philosophers believe that psychopaths are worthy of blame⁷ for their actions. For example, Matthew Talbert acknowledges that psychopaths are ‘morally blind’ (2012, p. 3). Nevertheless, he thinks, psychopaths are blameworthy because they are not generally immune to reasons. As Talbert notes, “the psychopath will often know when he has reason to do something and he can tell us what reasons he has” (p. 9). While he may not ‘have a grip’ on what others call moral reasons, “the psychopath can be characterized as holding that there is nothing to claims about moral status” (p. 11). This constitutes “express[ing] a morally significant perspective on the status of others” (p. 24) and the psychopath is therefore an “appropriate target of the judgments and attitudes that constitute moral blame” (p. 24).

Heidi Maibom (2008) insists on the responsibility of psychopaths on the grounds that they do not lack empathy altogether, and in addition possess many of the other moral capabilities that underlie judgments in the moral domain: “he does not lack empathy altogether. Furthermore, there are many other abilities that help give depth to ordinary people’s understanding of morality, none of which have been shown to be absent in the psychopath” (p. 173). For example, argues Maibom, psychopaths might

⁷ In this section, we will encounter a variety of morally laden terms such as blameworthy, responsible, accountable, etc. I am not changing them to create a unified terminology because I think it more important to preserve the authenticity of each account. Farther into the chapter, I will assess each position in the terms I have been using throughout this work.

refrain from evil deeds because they understand that acting on their impulses would deprive them of their autonomy. (p. 174) Furthermore, she argues that “there is little experimental evidence that psychopaths fail to appreciate what we like to call the higher things in life: religion, honor, purity, etc.” (p. 176), although she does admit that “[t]here is plenty of anecdotal evidence, however, that they are deficient in most, if not all, of these areas too.” (p. 176)

Even Robert D. Hare, the founder of the PCL-R, believes that psychopaths should be held accountable for their actions:

Still, some observers argue, psychopaths are deficient in the mental and emotional mechanisms needed to translate their knowledge of the rules into behavior acceptable to society... In my opinion, psychopaths certainly know enough about what they are doing to be held accountable for their actions. (p. 143)

Other philosophers think that the emotional deficits that make psychopaths unique exempt them from moral responsibility. One example of this position is Stephen Morse (2008), who does not believe that psychopaths are morally responsible for their actions:

psychopaths know the facts and the rules and are capable of manipulation of others to achieve their own ends, but they do not get the point of morality. It is as if they are color blind to moral concerns. The rights and interests of others have no purchase on their practical reasoning. Blaming and punishing such people is morally pointless, although it may be instrumentally warranted. (p. 208)

Cordelia Fine and Jeanette Kennett agree:

psychopaths are not accountable for their actions because while psychopathic offenders certainly appear to know that acts are prohibited by society or the law (and therefore know that their transgressions are

legally wrong), they do not appear to have the capacity to judge an act to be morally wrong. (2004, p. 432)

Paul Litton, finally, believes that psychopaths are not accountable for their actions because they do not have evaluative standards against which to assess and judge their own behavior:

It is not merely that the psychopath has not internalized any moral standards; he has a weakened capacity for possessing any kind of evaluative standards, and that weakened capacity is related to signs of irrationality... The key evidence for this conclusion is the psychopath's thorough immunity from feelings of regret, remorse, shame, and guilt. (pp. 375-378)

Other philosophers take an intermediate position, and suggest that psychopaths are somewhat morally responsible for their actions, but not fully so. The most sophisticated example of such a position is that of Walter Sinnott-Armstrong, who believes that the capacities necessary for moral responsibility come in degrees. Crucially, Sinnott-Armstrong takes this to be the case for individual actions — it is not merely the case that an agent might be responsible for some actions but not for others, but rather that they may be more or less responsible for individual actions (personal communication, 2015).

What is obvious from these various quotations is that the philosophers in question base their opinions on different facets of the psychopath's cognitive and emotional makeup: they regard different capacities as necessary or sufficient, respectively, for moral responsibility. It is to be expected that using different criteria to assess the moral responsibility of the psychopath yields divergent results. I hope that a

careful analysis of what it means to be a psychopath and to be responsible will resolve some of the controversy.

Misconceptions About Psychopathy

The first thing to be done in attempting to resolve the disagreement about the moral responsibility of the psychopath is to dispel some myths surrounding the concept. Psychopaths are often portrayed in one of two ways: One popular conception of psychopathy involves the ruthless, highly ambitious professional. The character of Gordon Gekko from Oliver Stone's movie *Wall Street* (1987) comes to mind, or — in the interest of gender equality — the demanding boss in David Frankel's 2006 movie *The Devil Wears Prada*. These depictions gained some academic acclaim in the work of Oxford research psychologist Kevin Dutton. In his book *The Wisdom of Psychopaths: What Saints, Spies and Serial Killers Can Teach Us About Success* (2012), Dutton identifies some of the traits associated with psychopathy (in particular, self-confidence, cold-heartedness, manipulateness, fearlessness, charm, cool under pressure, egocentricity, and carefreeness) and argues that they might, under certain conditions, prove beneficial to those who aspire to high-stress, high-responsibility, and high-prestige professional positions. On the basis of this contention, Dutton created the *Great British Psychopath Survey*, an online survey he ran from his personal website in 2011. Participants were asked to state their professions, and to answer a set of questions that attempted to assess the prevalence of the traits Dutton had identified as characteristic of psychopaths. Not particularly surprisingly, the survey found the highest levels of psychopathy among

CEOs, lawyers, media personae (particularly TV and radio), salespersons, and surgeons (Dutton, 2011). It is tempting to take these results to indicate that these professionals *are* actually psychopaths, and, moreover, that these portrayals are typical of psychopaths.

In direct opposition to the conception of psychopathy described in the previous paragraph is an equally familiar one that also garners significant attention in popular culture. This is the depiction of the psychopath as a crazed, often sadistic, serial killer. For the quintessential portrayal of this stereotype, think of the character of Norman Bates in Alfred Hitchcock's movie *Psycho* (1960). Consider, also, the portrayal of Buffalo Bill in Jonathan Demme's 1991 film *The Silence of the Lambs*. Both⁸ characters were inspired by the American killer Ed Gein, who murdered at least two women in the 1950s. Ed Gein also dug up and stole bodies from a local graveyard, and fashioned mementoes from their skin. Gein testified to being in a "daze-like" state while exhuming bodies, and recalled at least one occasion on which he emerged from this daze at the graveside, and returned home without fulfilling his mission. He also admitted that following the death of his mother, he began to fashion a suit from the skin of his female victims so that he could literally "crawl into the skin of his mother" (Ramsland, 2015, n. p.). After his arrest, Gein was initially found incompetent to stand trial, and sent to the Central State Hospital for the Criminally Insane in Waupun, Wisconsin. There, he was

⁸ For evidence that Norman Bates was, indeed, inspired by Ed Gein, follow this link: <http://darkecho.com/darkecho/horroronline/bloch.html/> Accessed March 24, 2015.

diagnosed with schizophrenia, but *not* with psychopathy. In 1968, Gein's doctors determined that he was sane enough to stand trial, and he was found guilty of first-degree murder. Nevertheless, in deference to his condition, he was found legally insane, and sent to the Mendota Mental Health Institute, where he died in 1984.

This conception of the psychopath as the equivalent of a crazed, out-of-his-mind killer is particularly explicit in online reactions to a 2014 study published in the *British Journal of Psychiatry* (Ando, Claridge, & Clark, 2014). The study attempts to correlate making a living as a stand-up comedian with four types of *psychotic* characteristics: belief in telepathy and paranormal behavior, cognitive disorganization, introverted anhedonia, and impulsive non-conformity. As such, the study attempts to relate the characteristics of comedians to those of individuals suffering from conditions such as schizophrenia and bi-polar disorder. The study, then, does not claim to reveal anything about psychopathic, as opposed to psychotic, tendencies in comedians. Nevertheless, a quick Google search of the terms 'comedians' and 'psychopathy' reveals that about half of the links on the first two pages represent this very study as establishing high levels of *psychopathy* among comedians. What this implies is that in popular culture, there is a strong tendency to conflate psychosis and psychopathy.

Neither of these popular portrayals of psychopathy adequately represent the actual condition: Although many ambitious professionals might exhibit some traits that are also characteristic of psychopathy, this does not mean that they exhibit these traits at sufficiently high levels to be classified as psychopaths. Neither does it mean that

possessing these traits, even at very high levels, is sufficient for a diagnosis of psychopathy. These people are not psychopaths — they are jerks. The deranged killer, by contrast, certainly has a mental disorder — but it is not psychopathy: he is *psychotic*, but that is a different disorder. Let us continue, then, by investigating what psychopathy is, rather than what it is not.

The PCL-R

The concept of psychopathy that is currently used most frequently in both clinical and forensic settings is based on The Psychopathy Checklist-Revised (PCL-R) as developed and revised by Dr. Robert D. Hare (1980; 1991; & 2003). Hare describes psychopathy as a condition that is characterized by a collection of twenty personality traits and recorded behaviors. He subdivides these items into two factors to distinguish between the items that represent the psychopath's interpersonal and affective deficits (Factor 1) and those that relate to his antisocial behavior and lifestyle (Factor 2). To these factors, he adds a third category of two items that assess the psychopath's sexual and marital habits.

Each of the two factors is itself subdivided into two facets: Factor 1 is divided into items that evaluate the psychopath's style of interaction with other people (Facet 1), and items that describe his affective personality (Facet 2), respectively. Factor 2 is separated into habits that relate to the lifestyle choices associated with psychopathy (Facet 3), and behaviors that exemplify the psychopath's tendency to behave anti-

socially and often in breach of the law (Facet 4). The following table represents the PCLR as divided into factors, facets, and other items:

Factor 1

Facet 1: Interpersonal

Glibness/superficial charm
 Grandiose sense of self-worth
 Pathological lying
 Cunning/manipulative

Facet 2: Affective

Lack of remorse or guilt
 Emotionally shallow
 Callous/lack of empathy
 Failure to accept responsibility for own actions

Factor 2

Facet 3: Lifestyle

Need for stimulation/proneness to boredom
 Parasitic lifestyle
 Lack of realistic, long-term goals
 Impulsivity
 Irresponsibility

Facet 4: Anti-social

Poor behavioral controls
 Early behavioral problems
 Juvenile delinquency
 Revocation of conditional release
 Criminal versatility

Other Items

Many short-term marital relationships
 Promiscuous sexual behavior

Most of the items on the checklist are self-explanatory, but some merit elaboration. This is the case, in particular, for items in Facet 2. These items concern the psychopath's emotional life, and are therefore inherently more difficult to assess from an external perspective, as compared to the items describing the psychopath's interactions with others, or his observable actions. 'Lack of remorse or guilt', according to the scoring instructions in the PCL-R manual,

describes an individual who shows a general lack of concern for the negative consequences that his actions, both criminal and noncriminal,

have on others. He is more concerned with the effects that his actions have upon himself than he is about any suffering experienced by his victims or damage done to society (1991, n. p.).

The psychopath might admit directly that he does not feel any remorse for his behavior, or he might claim that he does, but reveal in subsequent conversation that he does not, in fact, feel bad about what he has done. Conversely, an agent might simply deny that his actions have any consequences of concern to others. In extreme cases, the psychopath might actually feel that, rather than harming his victim, he has actually taught her a valuable lesson about life (Hare, 1993, p. 41). Even more startlingly, psychopaths frequently see themselves as the real victims of their actions (Hare, 1993, p. 43). In an interview with police inside Kingston Penitentiary in Ontario, Canada, a notorious killer, Paul Bernardo, laments what he considers the apparently undeserved isolation of his prison life, and the portrayal he receives in the media: "I sat here month after month after month and no one came... I turn on the TV in September, purely to clear the air, that Paul Bernardo was this crazy liar to police" (Benitah, 2008, para. 10).

'Emotionally shallow' refers to

an individual who appears unable to experience a normal range and depth of emotion. At times, he may impress as cold and unemotional. Displays of emotion generally are dramatic, shallow, short-lived; they leave careful observers with the impression that he is playacting and that little of real significance is going on below the surface (Hare, 1991, n.p.; 1993, p. 45).

It is thus not necessarily true that psychopaths have no emotions; rather, it is simply that their emotions lack both depth and persistence. In addition, many psychopaths show a diminished range of emotion. One emotion that is typically absent among the high-

scorers on this item is fear. Hare (1993) discusses a variety of studies that indicate that psychopaths do not experience the physiological responses that, in non-psychopaths, are associated with fear. A particularly telling experiment was conducted by David Lykken (2008), who measured the skin conductivity of subjects who knew that they would be receiving an electric shock in five seconds. In non-psychopathic participants, this knowledge cause anxiety, as measured by skin conduction sensors. The anxiety levels of psychopaths, by contrast, did not increase immediately upon being notified of the impending shock, and did not increase to the expected levels immediately before it occurred (2008). On the basis of such studies, Hare concludes that psychopathic fear is “incomplete, shallow, and largely cognitive in nature” (1993, p. 56). What is missing is the “emotional awareness of the consequences” (Hare, 1993, p. 54) of an action or situation. This emotional awareness is what steers non-psychopathic individuals away from particular courses of action. However, the psychopath who scores high on emotional shallowness does not appear to experience the emotional linkages between actions and consequences that would result in deterrence.

‘Callous/Lack of Empathy’ means, according to the PCL-R, that the psychopath has “a callous disregard for the feelings, rights, and welfare of others. He is only concerned with ‘Number 1’ and views others as objects to be manipulated. He is cynical and selfish. Any appreciation of the pain, anguish, or discomfort of others is merely abstract and intellectual” (1991, n. p.). In fact, many of the case studies presented by Hare (1993) and Cleckley (2011) suggest that psychopaths often find evidence of

weakness or pain distracting and annoying. For the psychopath, others' distress marks them as particularly easy targets, rather like blood in water does for sharks. It is this lack of interest in the "inner workings of 'loved' ones" (Hare, 1993, p. 45) that prevents psychopaths from maintaining ties with spouses and children, and from safeguarding the welfare of those individuals. Indeed, Hare suggests that it is this lack of empathy and resulting callousness that underlies some of the behaviors described by other items on the PCL-R. It is callousness, for example, that allows psychopaths to "parasitically bleed...other people of their possessions, savings, and dignity" and "aggressively do... and tak[e] what they want", as well as "engag[e] in an unending series of casual, impersonal, and trivial sexual relationships" (1993, p. 45).

'Failure to accept responsibility for own actions' indicates that the psychopath is "unable or unwilling to accept personal responsibility for his own actions (both criminal and noncriminal) or for the consequences of his actions. In excusing his behavior, he generally engages in rationalization, often by placing the blame on others (society, his family, accomplices, victims, the system, etc.). In extreme cases, "he may deny accusations made against him, despite overwhelming evidence" (PCL-R, n. p.; Hare, 1993, pp. 42-44). The quotation from Paul Bernardo given at the beginning of the present chapter nicely illustrates this point. For, although Bernardo broke a victim's arm, he indicates that he does not hold himself responsible for doing so — one has one's ups and downs. (Makin, 2008). Although Bernardo does not explicitly blame his victim for her injury in this particular interview, psychopaths often do just this. After all, the

psychopath tends to think, had the victim not struggled, he himself would not have been forced to react with so much violence.

The assessment of responses to items that address a subject's lifestyle and social interaction requires that the administrator familiarize herself with the individual's history. To do this, she must rely on external sources such as arrest records and testimony from family members. For this reason, the PCL-R applies only to agents eighteen years of age or more, as younger individuals simply do not possess a substantial enough history to enable the assessor to score the items in question. In addition to this requisite background research, administration of the PCL-R consists of a semi-structured interview that generally lasts from 1 to 4 hours. Over the course of this interview, a trained professional assessment administrator encourages the subject to share his life story, and steers the conversation toward experiences that will reveal to what extent the subject personifies each of the twenty characteristics that comprise the syndrome. For each of the items, the interviewer assigns a score of 0 (does not apply), 1 (applies somewhat), or 2 (applies) to the subject. The total score is calculated by adding the points awarded on each item. In North America, a clinical diagnosis of psychopathy necessitates a score of 30 or above, although some European countries employ the lower threshold of 26. Among those who score 30 on the PCL-R, it is customary among researchers in the field to isolate high-scoring psychopaths (34 and above) as a relevant group — these are the "true" psychopaths. It is generally agreed upon that there is something qualitatively different about individuals scoring 34 and above, when

compared to lower-scoring psychopaths (Kiehl, personal communication, November 2014).

Flavors of Psychopathy

True psychopaths are, fortunately, rare. This means that most of those who are diagnosed as such by means of the PCL-R “lose” points on some of the items. According to the PCL-R, then, psychopathy is a dimensional concept . This is by no means an idea with which all experts agree; according to Sinnott-Armstrong and Schaich Borg, there is still significant debate on whether psychopathy is, in fact, dimensional rather than categorical (2013, p. 108). However, they note that even those who challenge the dimensionality of psychopathy acknowledge recent empirical evidence suggesting, despite the practical need for a diagnostic cut-off point, that the disorder itself is indeed dimensional (2013, p. 108). On the basis of this evidence, we will treat the disorder as such.

Moreover, as the factor/facet model implies, psychopathy is a condition with a variety of possible presentations — in order to qualify as a psychopath, an individual must reach a total score of 30 on the PCL-R, but this total score might be made up of any constellation of individual scores on the 20 items. The PCL-R is thus equally consistent, for example, with A) a psychopath who scores extremely high on Facets 1 and 2, but shows only little impulsivity (we shall call him the controlled but uncaring psychopath, who is able to refrain from many of the anti-social behaviors in Facet 4); and B) a psychopath who scores high on Facet 4, and is highly irresponsible and impulsive, but

shows some traces of remorse for his anti-social actions (we shall call him the impulsive but caring psychopath). If both individuals obtain a total score of 30 or above on the PCL-R, they will receive the same diagnosis of psychopathy. Nevertheless, the condition will present very differently in each of these individuals.

The idea that psychopathy manifests itself in a number of ways is made explicit by Hervé (2007), who proposes four subtypes of psychopaths. His taxonomy distinguishes between the 'classic psychopaths,' who have the highest PCL-R scores overall and scores high on all facets, the 'explosive psychopaths,' who score relatively low on the interpersonal facet but comparatively high on the affective and lifestyle facets, the 'manipulative psychopaths,' who score high on the interpersonal and affective facets but relatively low on the lifestyle facet, and the 'pseudo-psychopaths,' who exhibit the behavioral characteristics but not the affective ones (Hervé, 2007). This taxonomy corresponds to typical classes of criminal activity: classic psychopaths and pseudo-psychopaths are more likely to have committed a variety of crimes in the course of their criminal career, and to defy authority. Pseudo and explosive psychopaths, meanwhile, exhibit the most destructive behavior to physical property — e.g. arson or vandalism — but the less fraudulent behavior, when contrasted to that of manipulative psychopaths. Overall, pseudo-psychopaths prove the least criminally problematic group, as pertains to severe interpersonal violence such as rape, murder, assault, and kidnapping (Hervé, 2007).

What these distinctions among types of psychopathic behavior and personality imply is that the question “Is the psychopath morally responsible for his actions” is formulated far too generally. Psychopathy affects (and, equally importantly, does not impact) individuals in a variety of ways. Thus, when we want to know whether a psychopath is an appropriate target for responsibility for some crime, we must know precisely how psychopathy affects him personally. In addition, we need to be specific about the nature of the crime under investigation. A particular psychopath’s impairment is likely to interfere with his capacity to refrain from committing some criminal acts, but not others.

Verdicts - Redux

It is, I believe, the complicated nature of the condition of psychopathy that underlies some of the controversy about the responsibility of psychopaths.

Matthew Talbert bases his opinion on the responsibility of psychopaths on the fact that they are not, as they are so often portrayed in popular culture, raving lunatics. He holds psychopaths responsible on the grounds that “the psychopath will often know when he has reason to do something and he can tell us what reasons he has” (p. 9). Psychopaths, then, are not impaired in the art of goal-attainment. When the psychopath decides that he has no reason to heed moral edicts, Talbert believes that he thereby expresses an opinion on the moral status of others — namely, that they don’t have one. This amounts to a positive and substantive affirmation that other individuals are not worthy of consideration. It is this active moral judgment that, according to Talbert,

makes the psychopath blameworthy for his actions. I find it questionable whether psychopaths actually make such substantive judgments about the worth of others — numerous quotations gleaned from the seminal books on psychopathy, such as *Without Conscience* (Hare, 1993) and *The Mask of Sanity* (Cleckley, 2011), imply that psychopaths simply do not give any thought at all to the moral worth of others. But if this is true, then it does not constitute an active judgment that other individuals do not have moral worth, but rather a complete lack of sensitivity to the question itself. Matthew Talbert's verdict on the responsibility of psychopaths, then, is based on a potential misunderstanding of the way psychopaths think.

Heidi Maibom (2009) holds psychopaths, as a group, morally responsible: she bases her verdict on the fact that *some* psychopaths are not entirely impaired in the area of empathy. This is undoubtedly true: as we have seen, the 20 features of psychopathy are all rated on a scale of 0, 1, or 2. It is therefore entirely possible to qualify as a psychopath — i.e. to score more than 30 — and still have *some* capacity for empathy. On my definition of accountability, this would make these psychopaths partially accountable: they might engage in evil deeds, but would still be capable of seeing *some* reason to abide by an empathy-based moral code. It must be pointed out, however, that “true” psychopaths who score 34 or above rarely score anything other than a 0 on the empathy factor, and thus Maibom's point does not apply to this group. Furthermore, Maibom points out that it has not been shown conclusively that psychopaths lack other abilities that might keep them from hurting others, such as the fear of losing their

autonomy by giving in to their impulses, or an appreciation of honor and duty. However, as the PCL-R shows, psychopaths are impulsive (and thus unlikely to consider such ephemeral things as autonomy before acting), and have shallow and variable commitments to values like honor or duty that are unlikely to be robust enough to prevent behavior that hurts others. Indeed, Maibom herself admits that there is plenty of anecdotal evidence that psychopaths are impaired in these areas. (2009, p. 176)

Maibom's assessment of the responsibility of psychopath thus applies only to a small subset of those who qualify as psychopaths — and that is, perhaps, as it should be. As indicated earlier in the chapter, psychopathy is a pathology that comes in many incarnations.

Robert D. Hare himself makes a similar point: although psychopaths might be mentally and emotionally impaired, they know enough about what is and isn't acceptable behavior in order to be held accountable. (p.143) Although I am restricting my discussion to moral rather than legal responsibility, Hare appears to suggest a quasi-legal criterion for moral responsibility: Observing other agents and learning from them, however theoretically, what is right and wrong is sufficient for moral responsibility. The quotations at the beginning of the chapter, as well as throughout it, have indicated that psychopaths often cannot discern what is acceptable behavior — but again, Hare's primary preoccupation with legal responsibility may set the bar lower than we do in this chapter.

On the side of philosophers who would exempt psychopaths from moral responsibility, Stephen Morse exempts psychopaths from moral responsibility for their actions on the grounds of their moral 'color blindness'. Unfortunately, he does not explain what 'moral blindness' means: if it means that the psychopath is unable to make moral judgments — like the blind person is unable to judge whether a shirt is red or blue without being told — then the analogy is invalid: psychopaths appear entirely able to make moral judgments; they simply don't follow them. However, if Stephen Morse takes 'moral blindness' to mean that the psychopath is completely oblivious to alternative courses of action that do not harm other people or considers himself to have no reason to follow them, then Morse's position is reminiscent of that of David Shoemaker, who insists on the importance of contrastive reason when assessing the responsibility of psychopaths. As promising as it sounds, Stephen Morse's verdict contains a critical amount of ambiguity.

Cordelia Fine and Jeanette Kennett exempt the psychopath from moral responsibility on the grounds that they are unable to make moral judgments (2004, p. 432). This contention has been drawn into question by Walter Sinnott-Armstrong and Jana Schaich Borg, who examined it in detail, and found that

neither psychopaths' lack of empathy nor their immoral behavior shows that psychopaths do not make moral judgments. Instead, psychopaths might make normal moral judgments but fail to translate this cognitive ability into normal emotion or motivations to avoid immoral actions. (2013, p. 111)

Schaich Borg and Sinnott-Armstrong present a plethora of evidence that the second interpretation is, in fact, the correct one. On the basis of their data, Schaich Borg and Sinnott-Armstrong conclude that “contrary to popular opinion, the current literature does not provide evidence [to support the] suggestion [that] psychopaths have severe moral cognitive defects” (p.108). Fine and Kennett’s judgment thus rests on evidence that has been discounted.

Paul Litton focuses on a different aspect of psychopathy. He claims that the psychopath

has a weakened capacity for possessing any kind of evaluative standards, and that weakened capacity is related to signs of irrationality... Individuals with psychopaths may not exhibit the signs of disordered thinking prevalent in persons with psychoses, but agents can be irrational in other ways. Young children are not responsible due to their lack of rational powers, and in many ways, persons with psychopathy are similar. (pp. 375–378).

David Shoemaker (2015) supports Litton’s point:

[W]hen [psychopaths] look through their own mind’s eye, it would seem that the evaluative landscape they see is just flat... They are unable to see (or have great difficulty seeing) facts about their own normative perspective in a ‘reasonish’ way... I am suggesting that psychopaths are evaluatively impaired generally. They tend not to have (or have poorly functioning) evaluative stances, stances from which various sorts of pursuits are viewed from the perspective of one *qua* persisting agent as worthwhile and are thereby transformed into prudential (guiding) ends. (p. 161–162).

Psychopaths simply, the quotation implies, simply don’t have evaluative standards — values that *really* mean anything to them. This is not to be taken to mean that psychopaths do not defend their commitments, however transient they might be,

with great bluster. Consider the example of notorious serial killer Paul Bernardo, who felt unfairly treated in prison: during an interview inside Kingston Penitentiary in Ontario, Canada, Bernardo laments what he considers the apparently undeserved isolation of his prison life, and the portrayal he receives in the media: “I sat here month after month after month and no one came... I turn on the TV in September, purely to clear the air, that Paul Bernardo was this crazy liar to police” (Benitah, 2008, para. 10). What Litton and Shoemaker imply, then, is that psychopaths are not responsible for his actions because nothing really matters to them, and if nothing matters, we cannot expect an agent to respect anything — or hold them responsible for violating a purported evaluative judgment that means nothing to him.

In this section, I hope to have demonstrated two things. Firstly, my intent was to show that the philosophers quoted at the beginning of the chapter employ very different standards in order to assess whether or not psychopaths are morally responsible. In addition, I hope to have shown that many of these philosophers base their verdicts on conceptions of psychopathy that are perhaps incomplete or, in some cases, have been demonstrated to be false. It is therefore no surprise that the opinions on the psychopath’s moral responsibility vary: in addition to using differing criteria for moral responsibility, the philosophers rely on conflicting conception of the syndrome.

Shoemaker and Psychopathic Responsibility

David Shoemaker’s opinion on the responsibility of psychopaths is particularly interesting because he employs an explicitly formulated and systematic conception of

what it means to be responsible for something. Shoemaker's treatment of psychopathy does not merely point to a particular facet of psychopathy as the deciding factor for moral responsibility, but rather applies his tripartite theory of responsibility to the question. Even so, Shoemaker is sometimes conflicted about what to say about the moral responsibility of the psychopath.

In 2011, he writes:

[i]s the psychopath attributability-responsible, that is, are his attitudes and action properly attributable to him? Yes, for they surely reflect his evaluative commitments, however shallow they may be.... is he answerability-responsible? Yes. Insofar as some of his actions and attitudes flow from evaluative commitments themselves grounded in reason, the psychopath may be an intelligible target of demands for justification.... is he accountability-responsible? I do not believe so. To be accountability-responsible is just to be susceptible for being the appropriate target of sanctions (...) communicating relationship-defining demand... But the key moral-relationship-defining demand — that you be sensitive to my interests — is simply lost on the psychopath. (2011, p. 628)

In 2015, Shoemaker still believes that psychopaths are not (fully) accountable for their actions, but now because they are unable to show the proper regard for others that is necessary for accountability: "the tripartite theory predicts that psychopaths are not accountable (or at least that their accountability is significantly mitigated) as they are impaired for identifying (...) regard." (p. 176) Identifying regard contrasts with detached or cognitive empathy: "detached empathizers are simply like spies in the target's head: While entirely painting their own set of cares and commitments, they are mere observers of the various psychological effects the world is having on their target." (p. 158) By contrast,

[r]egard instead is typically enabled or activated by identifying empathy, that is, genuine engagement with others. Rather than merely looking along the sight-line of another's mind's eye and then observing the psychological effects in her of what one has seen her see, when I engage in identifying empathy with someone, I take up the normative perspective of her mind's eye *as my own* in certain crucial respects..." (p. 158)

But "[i]f other people cannot matter (or if they matter very little) to psychopaths, then taking up their perspective — and regarding them generally — cannot matter either (or matters very little)." (p.160) The psychopath, then, is incapable of showing proper regard. As we saw in the chapter on accountability, Shoemaker sees regard as necessary for accountability. For these reasons, Shoemaker, in both 2011 and 2015, regards the psychopath as not accountable for his actions.

In the interest of fairness, it must be noted that while the above paragraphs stress the difference between Shoemaker's early and present positions on accountability, they are not all that different: in 2011, Shoemaker viewed "relationship-defining" demands as the class of actions for which agents were accountable — as opposed to actions which might make a relationship better, but do not define it (recall George and Martha). But even in 2011, Shoemaker defines the key relationship-defining demand that agents be sensitive to the interests of one another. This is not so very different from how he defines regard in 2015: to show regard is a "genuine engagement with others". Although the terminology is different, then, the point is similar.

By contrast, in 2011, Shoemaker is certain that the psychopath is answerability responsible: "[I]s he answerability-responsible? Yes. Insofar as some of his actions and

attitudes flow from evaluative commitments themselves grounded in reason, the psychopath may be an intelligible target of demands for justification.” (p. 628) This, according to Shoemaker, is indeed the case for psychopaths: “because some of his actions and attitudes flow from evaluative commitments themselves grounded in reason.” (p. 628) By 2015, Shoemaker is no longer so certain about the psychopath’s answerability status⁹:

The capacity rendering one eligible for answerability, recall, is the capacity for judgment about ‘instead of’ reasons, reasons in favor of doing something or having some attitude instead of some relevant contrary action or attitude. Now there seems no doubt that psychopaths have this capacity in many nonregard-based domains... [b]ut to the extent that my desire to avoid pain tends not even to appear to them to be a reason not to cause me pain, they are impaired in recognizing the relevant regard-based ‘instead of’ reason to which we would refer in the fleshed out version of our answerability demand... (p. 176)

Since psychopaths appear unable to entertain alternatives courses of action that do not cause pain to others as morally relevant, it is thus questionable, to Shoemaker, whether they are answerable in the moral domain.

I believe that Shoemaker’s vacillation reflects a lack of nuance in what it means to be answerable: psychopaths *are* answerable in certain senses, and they are *not* answerable in certain other senses. While my distinction among the senses of answerability thus cannot yield a unified answer, it can nevertheless shed light on the

⁹ Except for a very small class of action: Judging the worth of others in spite of knowing that they are bad judges thereof.

causes of Shoemaker's indecisiveness. It appears that in 2011, Shoemaker's concept of answerability depended on whether the psychopath could cite his *motivation*: the psychopath is answerable, Shoemaker appears to say, because his actions are motivated by 'evaluative judgments themselves grounded in reason' of which he is aware, and it would therefore be reasonable to ask the psychopath to cite the reasons for his actions. Certainly, on the surface, a psychopath can name the motivations for his actions.

Consider an example found in Kiehl:

When asked about his latest crime, he answered: 'Well, it's pretty simple really. This girl I was living with, well, she pushed all my buttons. I mean, she hit all three, right in a row and I just got pissed. I ran after her into the bathroom where she was drawing a bath and pushed her really hard into the wall... I mean, she was such a bitch to me that night.' (2015, pp. 26-27)

This psychopath certainly knows what motivated him, in the moment, to act as he did.

But in general, as we have seen, psychopaths have notoriously poor capacities of insight.

Kent Kiehl writes that "[the psychopath] was never going to be able to give a normal answer to the question of why... When [he] was asked why he killed those girls, his mind just went blank." (2015, p. 245) In addition, Kiehl claims that "sometimes psychopaths do things without reason, without motivation." (2015, p. 245) If Kiehl is right, the psychopath certainly has a diminished capacity for answering questions about his actions. Psychopaths, then, appear to be only partially able to provide interrogators with their motivations: their ability to do so depends on their level of introspection, and on the extent to which they act from motive rather than from mere impulse.

I cannot be entirely certain why Shoemaker abandoned this early conclusion, but the afore-discussed factors about the nature of psychopathy likely played a role. Firstly, as discussed above, Shoemaker came to agree with Paul Litton that psychopaths are ‘evaluationally impaired’ in all regards. It is thus questionable whether they indeed have ‘evaluative judgments grounded in reason’ robust enough to provide motivations for their actions. Moreover, even if the psychopath *had* such evaluative judgments, his ability to vocalize them is threatened by his well-established lack of introspective acumen: even if the psychopath *does* act on the basis of an underlying evaluative judgment, it is far from clear that he is able to introspectively identify these judgments. As Kent Kiehl says, “Psychopaths lack an ability to understand why they commit anti-social acts” (2014, p. 245). When one particular psychopath whom Kiehl interviewed was asked why he murdered his victims, “his mind just went blank” (2015, p. 245).

I suspect that these two reasons incited Shoemaker to shift his criterion for answerability to ‘instead of’ reasons — in order to be answerable for an action, an agent must be able to recognize morally relevant alternative courses of action as such. This criterion corresponds to my contention that psychopaths must be able to imagine what other people might do in similar situations, since there are usually other actions an agent might have carried out instead. When we want a full account from an agent about why anyone would do anything, what we want to know is why anyone would do *this* instead of *that*, where ‘that’ is what other people would regard as a morally relevant alternative. It is only in this way that we get a full and meaningful account of the range of reason

that the agent considers viable in a particular situation. As Shoemaker recognizes, “One could (...) have various local blindnesses, that is, be able to recognize the ‘instead of’ reasons only in some normative domains but not others. This would render one an answerable agent only within those former domains.” (2015, p. 82) In 2011, therefore, Shoemaker takes the answerability of psychopaths to be a question of motivation. In 2015, he takes it as a demand for personal answerability.

In 2015, Shoemaker is equally conflicted about the attributability responsibility of psychopaths. He acknowledges that in the past he thought it perfectly acceptable to think of psychopaths with disdain and to label them “cruel,” “manipulative”, “nasty”, and the like. (2011) However, Shoemaker comes to wonder how,

if psychopaths have impaired capacities for caring about anything *and* they have impaired capacities for an evaluative stance (given their deficits of even self-regard), then in what sense could they be said to have a deep self at all for such attitudes to express? (p. 179)

He elaborates

if psychopaths indeed tend not to have a general evaluative stance, and if they tend not to have genuine emotional dispositions either, it would seem to follow that they would tend not to have cares, commitments, or care-commitment clusters that I have argued constitute the deep self, so nothing could be attributable to them. (p. 180)

This is reminiscent of the worry of Paul Litton, who does not believe that psychopaths operate from a strong deep self because they have “a weakened capacity for possessing any kind of evaluative standards.” (pp. 375–378) But not having a deep self, according to most traditional deep self theories, makes an agent not attributability responsible.

Despite his doubts, however, Shoemaker claims that “[a]t the end of the day, I think

there is a quite plausible story that can be told about how psychopaths are fully attributability-responsible: they *are* cruel, manipulative, etc.” (2015, p. 182)

Psychopathic Responsibility, Susanne Style

Shoemaker, then, is conflicted about the psychopath’s responsibility status when it comes to attributability responsibility and answerability responsibility, but not when it comes to accountability responsibility. There is, therefore, not only a conflict about the responsibility of psychopaths among philosophers, but even within the writings of the single philosophers who has engaged with the topic at the greatest length. I think that my proposed definitions of the different types of answerability can help dispel the conflict intra-Shoemaker, and inter-philosopher.

I will begin with the concept of attributability. Shoemaker, as quoted in the previous section, is not certain whether psychopaths qualify for attributability responsibility because the lack of depth in their cares and commitments (even to themselves) means that they do not have a robust deep self. I grant Shoemaker and Litton’s worry that psychopaths have ‘thin’ deep selves: whereas the rest of us have relatively stable and usually varied commitments and cares, in the psychopath these commitment are fleeting and superficial. But there are, as both Shoemaker and the PCL-R remind us, personality traits that *are* characteristic of psychopaths: they are cruel, manipulative, etc. I would like to remind the reader of my highly fictionalized account of Vlad the Impaler from an earlier chapter: my conclusion, on that occasion, was that we could not attribute Vlad’s *particular* choice of torture method to him because it is, in a

sense, arbitrary: he could have liked to keep his prisoners in dungeons instead of impaling them. Nevertheless, his actions reflect a deep and stable cruelty as part of his deep self. So it is with the psychopath: we may not be able to attribute to him specific actions and commitments because they are so fleeting, but these actions are reflections of his deep-seated callousness. It is true, then, that psychopaths have very impoverished deep selves, but those impoverished deep selves are — in my opinion — no less expressed in their actions than are those of non-psychopaths, and are thus attributable to them.

My verdict on the answerability of psychopaths is more complicated. Whether or not psychopaths are appropriate targets for demands of interpersonal justifiability is a complex question. As Sinnott-Armstrong and Schaich Borg (2013) demonstrate, intelligent and observant psychopaths have the capabilities of learning to make moral judgments. While these quick-witted psychopaths do not, themselves, see certain actions as violating their personal moral code, they have learned what types of actions violate the moral codes of those around them, and in what ways. There is nothing from keeping a well-trained psychopath from watching footage of children being abused and exclaiming “Gee, this is terrible!” without actually feeling anything of the emotions here highlights. Robert D. Hare illustrates the concept in his classic book *Without Conscience*: “Often, he was able to talk his way out of trouble. ‘I’d just look them straight in the eye and feed them shit. It was great. I still do it. My mother bought it for a long time’” (1993, p. 162).

The problem is, however, that eventually, even the ‘shit’-producing psychopath’s mother stopped buying it. This is, I suggest, because real-life moral conversations are often more involved than the short-answer format typically employed in scientific studies: when psychopaths are actually challenged to elaborate on their own moral reasoning, they often say very strange things. (This is, partially, the reason why the PCL-R consists of an extended semi-structured interview rather than a series of individually contained questions — we want to see what the psychopath says beyond some potentially well-rehearsed standard lines.) Recall, for example, the two quotations from Arias and Bernardo at the beginning of this chapter. Consider a further example from *Without Conscience*:

Do I care about other people? That’s a tough one. But, yeah, I guess I really do... but I don’t let my feelings get in the way... Do I feel bad if I have to hurt someone? Yeah, sometimes. But mostly it’s like... uh... [laughs]... how did you feel the last time you squashed a bug? (quoted in Hare, 1993, p. 33).

These examples imply, I contend, that psychopaths can indeed answer a set of standard moral questions. They can even, if they are smart, expand upon this set of answers to form a rather impressive repertoire of moral-sounding platitudes. Hervey Cleckley writes that the psychopath

cannot be taught awareness of significance of what he fails to feel. He can be taught to use the ordinary words and, if he is very clever, even extraordinarily vivid and eloquent words which signify these matters to other people. He will also learn to reproduce all the pantomime of feeling; but as Sherrington said of the decerebrate animal, the feeling itself does not come to pass. (1976, p. 374)

However, as the extrapolation from this base set of moral-ish proclamations gets more involved or extended, psychopaths eventually betray themselves. At a certain level of complexity, the evidence appears to show, coherent moral discourse is no longer possible. This claim is not mere contention: as the work by Simon, Holzenberg and Unger (1951) illustrates, females who scored high on psychopathic traits provided deviant answers to a moral dilemma questionnaire when this questionnaire was in free form, but not when it was in multiple choice format. They write that

the psychopathic girl has learned social values and is able to recognize them where they are overtly presented (multiple choice test) but where the psychopath is thrown upon her own resources resolving a situation (completion test), these learned values are not readily available as guides, apparently because of their superficial incorporation into the personal structure" (pp. 146–147).

Psychopaths, then, are subject to demands for interpersonal answerability commensurate with their powers of observation, their desire to learn from others in order to blend in (peers, lawyers, etc.), and the depth of the answer that is being sought.

Are psychopaths personally answerable? Ample evidence from the PCL-R indicates that psychopaths possess the type of moral 'blind spot' Shoemaker identifies where harm to others is concerned: it is simply not a factor in their deliberations. What this means is that because of this particular blind spot, psychopaths are inappropriate targets for demands for *personal answerability* regarding actions that cause harm to others: those actions do not properly violate the psychopath's personal code of conduct because his code of conduct, insofar as he has one, does not extend to the actions in question. This moral blind spot also makes it pointless to ask psychopaths to imagine

why other people would act differently from how they themselves act: they cannot imagine, in the harm-based moral domain, why *anyone* would take into consideration purported moral edicts (except for practical reasons). It might, however, remain reasonable to ask the psychopath to cite his motivations: sometimes they can give us (very superficial) reasons for their actions, but often they simply lack the insight or desire to ruminate and on own motivations.

The point is, then, that the answerability status depends on how we interpret the questions “Why did you do *x*?”. Are psychopaths personally answerable in the moral domain? No, because they do not have a personal code of conduct that guides their behavior in that domain. Are psychopaths interpersonally answerable? Only to the extent to which they are clever and observant, and the questions are kept general and short. As discussed in the chapter on answerability, it appears to be *interpersonal justifiability* that ultimately best captures Shoemaker’s concept of answerability. I would therefore have to amend his conclusion that psychopaths are not answerability responsible to state that they are somewhat answerability responsible in the ways that seem to matter to Shoemaker, but not fully so.

Accountability

The most important question — legally, practically, and medically — is whether psychopaths are morally accountable for their actions. Recall that Shoemaker does not think so on the grounds that “they are impaired for identifying (...) regard.” (p. 176). I agree with his conclusion, but for different reasons: according to my definition of

accountability, an agent is accountable for an action if they violate an authority's moral code of conduct, *despite* their ability to regard that code of conduct as a reason for action.

Much of the moral code of conduct that exists among members of most, if not all, societies depends on the avoidance of gratuitously causing harm to one another. There is ample evidence that psychopaths do not regard such harm-avoidance-based code of conduct as reason for action — this means, in other words, that psychopaths do not regard cooperation among individuals, and the harm avoidance it entails, as binding *for them*. Robert D. Hare puts the point succinctly:

Psychopaths view people as little more than objects to be used for their own gratification. The weak and the vulnerable — whom they mock, rather than pity — are favorite targets. 'There is no such thing, in the psychopathic universe, as the merely weak,' wrote psychologist Robert Rieber. 'Whoever is weak is also a sucker; that is, someone who demands to be exploited.'... [Psychopaths] are indifferent to the rights and sufferings of family members and strangers alike. (pp. 44–45)

Hare elaborates: "Psychopaths don't have an inner voice to guide them; they know the rules but follow only those they choose to follow, no matter what the repercussions for others." (pp. 75–76).

This conclusion dovetails with that of Sinnott-Armstrong and Schaich Borg (2013), who believe that psychopaths do indeed make moral judgments — but then don't see a reason to follow through. In other words, they understand what moral authority demands; but decide that they need not heed it. This does not mean that psychopaths cannot appreciate *any* reason to comply with common morality. Take, for example, this quotation from Ted Bundy: "I quickly discovered that the greatest obstacle

to my freedom, the greatest block and limitation to it, consists in the insupportable 'value judgment' that I was bound to respect the rights of others." (quoted in Polman & Friesen, 2011, p. 30) Ted Bundy, then, was smart enough to see *some* reason for following the moral code of his society — it just wasn't the right kind of reason. What guided him (ultimately not well enough) was self-interest, not the content of the moral code of those around him. While Bundy was cognitively nimble enough to know the contents of that moral code (and was therefore interpersonally answerable), he did not regard it as a reason for action. On my account, he is therefore not morally accountability responsible for his actions. (Legal accountability is, of course, a different question). As Stephen Morse puts it: "Blaming and punishing such people is morally pointless, although it may be instrumentally warranted." (p. 208)

Conclusion

What I hope to have demonstrated in this dissertation is that it is possible for an agent to be morally responsible in some ways, and not in others. The case of the psychopath illustrates this. Which sense of moral responsibility is relevant to us depends largely on our interests in a particular context. If we want to know what sort of a person an agent is, we look toward attributability. When we want to know why an agent acts the way he does, answerability is the right sense of moral responsibility. When we want to know how to react to that agent, how to treat him, accountability provides the answer.

As final note, I would like to stress that the focus of this thesis has been the *moral* responsibility of agents. What is probably at least as important is the *legal* responsibility of these agents. I do not feel qualified to take on this question because, for one, the criteria for legal responsibility are so different from the moral requirements that the two issues deserve different dissertations. Secondly, I do not know enough about the legal requirements for responsibility to take on this latter question. Were I to continue this project, my next step would be to inform myself on these requirements by taking the relevant courses at either UNC or at Duke, in preparation of writing a companion piece to the current dissertation that considers the issue of psychopaths' responsibility from a legal perspective.

Works Cited

- Aharoni, E., Sinnott-Armstrong, W., & Kiehl, K. (2012). Can psychopathic offenders discern moral wrongs? A new look at the moral/conventional distinction. *Journal of Abnormal Psychology, 121*(2), 484-497.
- Aharoni, E., Sinnott-Armstrong, W., & Kiehl, K. (2014). What's wrong? Moral understanding in psychopathic offenders. *Journal of Research in Personality, 53*, 175- 181.
- Ando, V., Claridge, G., & Clark, K. (2014). Psychotic Traits in Comedians. *The British Journal of Psychiatry, 204*(5), 341-345.
- Arpaly, N., & Schroeder, T. (1999). Praise, Blame and the Whole Self. *Philosophical Studies 93*(2): 161–188.
- Benitah, S. (2008). Bernardo police interview released to media. *CTV News Toronto*. 10 June 2008.
- Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition, 57*, 1-29.
- Blair R. J. R. (2008). The cognitive neuroscience of psychopathy and implications for judgments of responsibility. *Neuroethics, 1*, 149-157.
- Blair, J., Mitchell, D., & Blair, K. (2005). *The Psychopath: Emotion and the Brain*. Blackwell Publishers.
- Brink, D. (1996). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Caldwell, M., Skeem, J., Salekin, R., & Van Rybroek, G. (2006). Treatment Response of Adolescent Offenders With Psychopathy Features : A 2-Year Follow-Up. *Criminal Justice and Behavior 33*(5), 571-596.
- Camus, A. (1967). *The Stranger*. London: Penguin Modern Classics.
- Chivers, T. (2014). Psychopaths: can you spot one? *The Telegraph*, April 6, 2014.
- Cleckley, H. (2011). *The Mask of Sanity: An Attempt to Clarify Some Issues about the So-Called Psychopathic Personality (3rd ed)*. Whitefish: Literary Licensing.
- Duff, A. (1977). Psychopathy and Moral Understanding. *American Philosophical Quarterly, 14*.

- Dutton, K. (2012). *The Wisdom of Psychopaths: What Saints, Spies and Serial Killers Can Teach Us About Success*. New York: Scientific American.
- Dwyer, S. (1999). Moral Competence. In K. Murasugi and R. Stainton (Eds.), *Philosophy and Linguistics*. Westview Press, 169-190.
- Dwyer, S. (2000). The mad, the bad, and the psychopath. *The Philosopher's Magazine* 9(9), 32-33.
- Einstein field equations - Wikipedia, the free encyclopedia. (n.d.). Retrieved November 4, 2015, from https://en.wikipedia.org/wiki/Einstein_field_equations
- Engel, S., et al. (2005). Impulsivity and compulsivity in bulimia nervosa. *International Journal of Eating Disorders* 38(3): 244–51.
- Fine, C, & Kennett, J. (2004). Mental impairment, moral understanding and criminal responsibility: Psychopathy and the purposes of punishment. *International journal of law and psychiatry*, 27.
- Finerman, W. (Producer), & Frankel, D. (Director). (2006). *The devil wears prada* [Motion picture]. USA: Fox 2000.
- Fingarette, H. (1988). *Heavy Drinking: The Myth of Alcoholism as a Disease*. Berkeley: University of California Press.
- Foot, P. (1978). The problem of abortion and the doctrine of the double effect. In *Virtues and Vices and Other Essays in Moral Philosophy* (19-32). Berkeley and Los Angeles: University of California Press.
- Forth, A., Bo, S., & Kongsersley, M. (2013). Assessment of Psychopathy: The Hare Psychopathy Checklist Measures. In K. Kiehl and W. Sinnott-Armstrong (Eds.), *Handbook on Psychopathy and Law* (5-33). New York: Oxford University Press.
- Frankfurt, H. (1969). Alternate Possibilities and Moral Responsibility, *Journal of Philosophy*, 66: 829–39.
- Frankfurt, H. (1971). Freedom of the Will and the Concept of a Person, *Journal of Philosophy*, 68: 5–20.
- Gendler, T. S. (2008). *Journal of Philosophy* 105(10), 634-663.
- Glannon, W. (2008). Moral Responsibility and the Psychopath. *Neuroethics* 1, 158-166.
- Glenn, A., Koleva, S., Iyer, R., Graham, J., & Ditto, P.H. (2010). Moral identity in psychopathy. In *Judgment and Decision Making*, 5, 497-505.

- Greene J.D., et al. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105-2108.
- Greene J.D. (2007). Why are VMPFC patients more utilitarian?: a dual-process theory of moral judgment explains. *Trends in Cognitive Science* 11, 322-323.
- Greene, J. D. (2008). The Secret Joke of Kant's Soul. In W. Sinnott Armstrong (Ed.), *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development*. Cambridge, MA: MIT Press.
- Greene, J.D (2013). *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. London: Penguin Press.
- Greene, J.D. (2015). The rise of moral cognition. *Cognition*, 135, 39-42.
- Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Difference*, 1(2): 111-120.
- Hare, R. D. (1991). *Manual for the Hare Psychopathy Checklist - Revised*. Toronto, ON: Multi-Health Systems, Inc.
- Hare, R. D. (1993). *Without Conscience: The Disturbing World of the Psychopaths Among Us*. New York: The Guildford Press.
- Hare, R. D. (2003). *Manual for the Hare Psychopathy Checklist - Revised (2nd ed)*. Toronto, ON: Multi-Health Systems.
- Hare, R. M. (1952). *The Language of Morals*. Oxford: Clarendon Press.
- Harenski, C., & Kiehl, K. (2011). Emotion and Morality in Psychopathy and Paraphilias. *Emotion Review*, 3(3), 299-301.
- Hervé, H. (2007). Psychopathic Subtypes: Historical and Contemporary Perspectives. In H. Hervé and J. Yuille (Eds.), *The Psychopath: Theory, Research, and Practice* (431-460). Mahwah: Lawrence Erlbaum.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814-834.
- Haidt, J., & Graham, J. (2007). When Morality Opposes Justice: Conservatives Have Moral Intuitions that Liberals may not Recognize. *Social Justice Research* 20, 98-116.
- Hiatt, K. D., Schmitt, W.A., & Newman, J. P. (2004). Stroop tasks reveal abnormal selective attention among psychopathic offenders. *Neuropsychology* 18: 50-9.

- Hitchcock, A. (Producer), & Hitchcock, A. (Director). (1960). *Psycho* [Motion picture]. USA: Shamley Productions.
- Hobbes, T., & Gaskin, J. C. A. (1998). *Leviathan*. Oxford: Oxford University Press. Hume D. (1978). *A Treatise of Human Nature*. Oxford University Press.
- Johns, J. H., & Quay, H. C. (1962). The effect of social reward on verbal conditioning of psychopathic and neurotic military offenders. *Journal of Consulting and Clinical Psychology*, 26: 217-220.
- Kennett, J., & Fine, C. (2008). Internalism and the Evidence from Psychopathy and Acquired Sociopathy. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development*. Cambridge: MIT Press.
- Kiehl, K. (2014). *The Psychopath Whisperer: The Science of Those Without Conscience*. New York: Crown Publishers.
- Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, 7, 708-714.
- Kohlberg, Lawrence (1981). *Essays on Moral Development, Vol. I: The Philosophy of Moral Development*. San Francisco, CA: Harper & Row.
- Korsgaard, C. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Levy, N. (2007). The responsibility of the psychopath revisited. *Philosophy, Psychiatry, and Psychology*, 14(2), 129-138.
- Litton, P. (2014). Criminal Responsibility and Psychopaths: Do Psychopaths Have a Right to Excuse? In Ki. Kiehl and W. Sinnott-Armstrong (Eds.), *Handbook on Psychopathy and Law* (275-296). New York: Oxford University Press.
- Lykken, D. (2008) The Psychopath and the Lie Detector. *Psychophysiology* 15(2).
- Maibom, H. (2005) Moral Unreason: The Case of the Psychopath. *Mind and Language* 20(2).
- Maibom, L. (2009) Feeling for Others: Empathy, Sympathy, and Morality. *Inquiry*, 52(5), 483-499.
- Makin, K. (2008). Anxiety over sex drove him, Bernardo told police. *The Globe and Mail*. 21 June, 2008.

- McDowell, J. (1978). Are Moral Requirements Hypothetical Imperatives? *Aristotelian Society Supplementary Iiii*, 13-29.
- Morse, Stephen J. (2008) Psychopathy and criminal responsibility. *Neuroethics*, 1(3), 205-212.
- Nadelhoffer, T., & Sinnott-Armstrong, W. (2012). Neurolaw and Neuroprediction: Potential Promises and Perils. *Philosophy Compass*, 7(9), 631-642.
- Nagel, T. (1970). *The Possibility of Altruism*. Oxford: Oxford University Press.
- Newman, J., Patterson, C., Howland, E., & Nichols, S. (1990). Passive Avoidance in Psychopaths: The Effects of Reward. *Personality and Individual Differences* 11:1101-1114.
- Nichols, S. (2002). How Psychopaths Threaten Moral Rationalism. *The Monist* 85(2), 285-304.
- Paxton, J., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in Cognitive Science*, 2, 511-527.
- Piaget, Jean (1932). *The Moral Judgment of the Child*. London: Kegan Paul, Trench, Trubner and Co.
- Pojmna, L, & Frieser, J. (2011). *Ethics: Discovering Right and Wrong*. Wadsworth.
- Ravizza, M., and Fischer, J.M. (1998). Responsibility and Control: A Theory of Moral Responsibility. Cambridge: Cambridge University Press.
- Prinz, J. (2007). *The Emotional Construction of Morals*. Oxford: Oxford University Press.
- Ramsland, Katherine (2012). A True Necrophile. *Crime Library*. Retrieved from www.crimelibrary.com on Jan. 27, 2015.
- Rest, J., Cooper, D., Coder, R., Masanz, J., & Anderson, D. (1974). Judging the important issues in moral dilemmas — An objective measure of development. *Developmental Psychology*, 10(4), 491-501.
- Robinson, P. H., & Kurzban, R. (2007). Concordance & conflict in institutions of justice. *Minnesota Law Review*, 91, 1829-1907.
- Roskies, A. (2003). Are Ethical Judgments Intrinsically Motivational? Lessons from 'Acquired Sociopathy'. *Philosophical Psychology* 16(1), 51-66.

- Scanlon, T. M. (1998). *What We Owe to Each Other*. Cambridge, Mass.: Harvard University Press.
- Schaich Borg, J. (2013). Subcomponents of Psychopathy have Opposing Correlations with Punishment Judgments. *Journal of Personal Social Psychology*, 105(4), 667–687.
- Schaich Borg, J., & Sinnott-Armstrong, W. (2013). Do Psychopaths Make Moral Judgments? In K. Kiehl and W. Sinnott-Armstrong (Eds.), *Handbook on Psychopathy and Law* (107-130). New York: Oxford University Press.
- Shoemaker, D. (2011). Attributability, Answerability, and Accountability: Toward a Wider Theory of Responsibility. *Ethics*, 121, 602-632.
- Shoemaker, D., & Faraci, D. (2014). Huck vs. JoJo: The Moral Ignorance (A)symmetry. In S. Nichols, J. Knobe, and T. Lambrozo (Eds.), *Oxford Studies in Experimental Philosophy*, v. 1. Oxford: Oxford University Press, 7-27.
- Shoemaker, D. (2015). *Responsibility from the Margins*. Toronto: Oxford University Press.
- Sinnott-Armstrong, W., & Pickard, H. (2013). What is addiction? In K. W. M. Fulford et al. (Eds.), *The Oxford Handbook of Philosophy and Psychiatry*. Oxford: Oxford University Press.
- Sinnott-Armstrong, W. (2014). Do Psychopaths Threaten Internalism? In Schramme, T. (Ed.), *Being Amoral: Psychopathy and Moral Incapacity*. Cambridge: MIT Press.
- Smith, A. (2005). Responsibility for Attitudes: Activity and Passivity in Mental Life. *Ethics* 115, 236-271.
- Smith, A. (2008). Control, responsibility, and moral assessment. *Philosophical Studies* 138, 367-392.
- Smith, A. (2012). Attributability, Answerability, and Accountability: In Defense of a Unified Account. *Ethics* 122, 575-589.
- Sripada, C. S. (2008). Nativism and moral psychology: Three models of the innate structure that shapes the contents of moral norms. *Moral Psychology*, 1, 319–343.
- Sripada, C. S. (n.d.). Self-expression: A Deep Self Theory of Moral Responsibility. *Published Online: 12 August 2015*.
- Sripada, C. S. Frankfurt's Willing and Unwilling Addict. Unpublished.
- Stone, O. (Producer), & Stone, O. (Director). (1987). *Wall street* [Motion picture]. USA: 20th Century Fox.

- Strawson, P. F. (1962). Freedom and Resentment. *Proceedings of the British Academy*, 48, 1–25.
- Talbert, M. (2008). Blame and Responsiveness to Moral Reasons: Are Psychopaths Blameworthy? *Pacific Philosophical Quarterly* 89, 516-35.
- Talbert, M. (2012). Accountability, Aliens, and Psychopaths: A Reply to Shoemaker. *Ethics*, 122(3), 562-574.
- Turiel, Elliot (1983). *The Development of Social Knowledge: Morality and Convention*. Cambridge: Cambridge University Press.
- Utt, Kenneth, Edward Saxon, Ron Bozman, (Producers), & Demme, J. (Director). (1991). *Silence of the lambs* [Motion picture]. USA: Orion Pictures.
- Vargas, M., & Nichols, S. (2007). Psychopaths and Moral Knowledge. *Philosophy, Psychiatry, & Psychology*, 14(2), 157-162.
- Wallace, R.J. (1999). Three Conceptions of Rational Agency. *Ethical Theory and Moral Practice* 2(3): 217–242.
- Watson, G. (1996). Two Faces of Responsibility. *Philosophical Topics*, 24, 227–48.
- Wolf, S. (1990). *Freedom within Reason*, Oxford: Oxford University Press.
- Young, L., Koenigs, M., Kruepke, M., & Newman, J. P. (2012). Psychopathy Increases Perceived Moral Permissibility of Accidents. *Journal of Abnormal Psychology*, 121(3), 659-667.

Biography

Susanne Meyer was born on June 17, 1977, in Zurich, Switzerland. Even during her high school years, she showed a keen interest in philosophy. In 1997, after graduating from Kantonsschule Oerlikon in Zurich, Susanne — in search of new experiences — moved to Canada, where she acquired a BA from St. Mary's University in Halifax, NS. She majored in Philosophy, and minored in French.

Starting in 2001, she studied for an MA in Philosophy at Simon Fraser University in Burnaby, BC, Canada, and completed a Master's thesis entitled *Thinking What We Want: The Variety and Nature of Unintended Thought*. After a brief hiatus from philosophy, Susanne began her PhD studies at the University of Toronto in 2005.

However, after one semester in that program, her personal circumstances altered, and she relocated to the United States to join her husband. In 2007, Susanne decided to continue her studies, and applied to Duke Philosophy PhD program. She was accepted, and started in the fall of 2008. Her continuation through the program was slowed down considerably by the arrival of three children in short order, and by a significant health crisis in Susanne's immediate family. Nevertheless, in November of 2016, she finally submitted her dissertation.