

# **Towards a Robust Theory of Evil Character**

By

Aaron Richter

A Thesis Submitted to  
Saint Mary's University, Halifax, Nova Scotia  
In Partial Fulfillment of the Requirements for  
the Degree of Masters of Arts in Philosophy

December 7, 2015, Halifax, Nova Scotia

Copyright Aaron Richter

Approved: Dr. Todd Calder

Supervisor

Approved: Dr. Mark Mercer

External Examiner

Approved: Dr. Sheldon Wein

Reader

Date: December 7, 2015

## **Abstract**

### Towards a Robust Theory of Evil

by Aaron Richter

December 7, 2015

Abstract: This thesis demonstrates that current theories of evil character, specifically those asserted by John Kekes, Colin McGinn, Daniel Haybron, Peter Brian Barry, and Luke Russell, are not robust accounts of evil character. A robust account of evil character provides intuitively acceptable necessary and sufficient conditions for evil character, creates a clear distinction between bad and evil characters, and includes a conception of degrees of evil. After examining problems with extant accounts of evil character, I argue that the locus of evil character lies not in our feelings, desires, actions, or dispositions, but in the conclusions we draw about what is morally acceptable.

## **Acknowledgements**

I'd like to thank heartily Bethany Daigle (University of New Brunswick) without her unceasing devotion and friendship this project could not have been completed. Dr. Todd Calder (Saint Mary's University) and Dr. Shelagh Crooks (Saint Mary's University) also have my deepest gratitude for inspiring and nurturing my love of philosophy from the very beginning. Finally, to my grandmother, for all the countless years of encouragement and support, I dedicate this thesis; may you rest in peace.

## Table of Contents

<b>Acknowledgements</b> .....	i
<b>Table of Contents</b> .....	ii
<b>Chapter 1: Introduction</b> .....	1
<b>Chapter 2: The Perverse Benefactor</b> .....	5
2.1 A New Example .....	5
2.2 Justifying the Perverse Benefactor .....	6
2.3 The Characteristics of a Perverse Benefactor .....	8
2.4 Criticism against Good Desires.....	14
2.5 Summary .....	15
<b>Chapter 3: John Kekes’s Frequent Evil-doing Account</b> .....	17
3.1 The Account .....	17
3.2 How Frequently is Enough .....	20
3.3 The Malicious Quadriplegic.....	20
3.4 The Specific Motivation Criticism .....	21
3.5 Summary .....	22
<b>Chapter 4: Colin McGinn’s Affective Account</b> .....	25
4.1 The Account .....	25
4.2 Excluding Motivation .....	26
4.3 Frequency and Intensity.....	27
4.4 Evil without Reversed Disposition.....	28
4.5 Summary .....	29
<b>Chapter 5: Daniel Haybron’s Consistency Account</b> .....	31
5.1 The Account .....	31
5.2 An Unclear Condition .....	34
5.3 Lacking Descriptive Power .....	34
5.4 No Degrees of Evil .....	35
5.5 Evil Characters that are Moved by the Good.....	37
5.6 Summary .....	38

<b>Chapter 6: Peter Brian Barry’s Extremity of Vice Account</b> .....	40
6.1 The Account .....	40
6.2 Lacking Degrees of Evil.....	42
6.3 Too Narrow .....	44
6.4 Summary .....	45
<b>Chapter 7: Luke Russell’s Dispositional Account</b> .....	47
7.1 The Account .....	47
7.2 Unclear Elements .....	50
7.3 The Lazy.....	52
7.4 Summary .....	54
<b>Chapter 8: Towards a Robust Account</b> .....	57
8.1 Chapter Outline.....	57
8.2 The Five Accounts .....	57
8.3 A Robust Sketch .....	61
<b>Bibliography</b> .....	68

## **Chapter 1: Introduction**

This project's purpose is to demonstrate that the theories of evil character asserted by John Kekes, Colin McGinn, Daniel Haybron, Peter Brian Barry, and Luke Russell are not robust accounts of evil character. A robust account of evil character is one that is intuitively plausible, creates a clear distinction between bad and evil characters, and includes a conception of degrees of evil.

Our intuitions about evil are the vague, yet undeniable, standard to which we must conform any theory of evil character. Without our intuitions about evil guiding us, we are incapable of developing an account that explains evil character correctly. Therefore, the first criterion that our prospective accounts must meet is intuitive plausibility. An intuitively plausible account is one that provides plausible necessary and sufficient conditions for evil character to obtain and is capable of accommodating all possible types of evil character. An account is not intuitively plausible if it leaves a type of evil character unaccounted for, or if it provides false necessary or sufficient conditions for a person to be evil.

A robust account of evil character must also distinguish between the very bad person and the evil person. Although all three criteria arise directly from our intuitions about evil, this distinction is important enough to warrant its own listing as a criterion. We moral beings tend to think that not every bad person is evil. Therefore, it is desirable for a robust account of evil character to reflect that intuition. A robust account must

establish parameters that make certain that only the very worst sorts of people qualify as evil without mistakenly apprehending the merely very bad as evil.

Similarly, and finally, a robust account of evil character must address the concept of degrees of evil. I will demonstrate in this project that the concept of degrees of evil is inextricable from our intuitions about evil character. The position that there are, or could be, degrees of evil bothers many of the authors whose accounts we will examine. We will demonstrate that each author presupposes a concept of 'degrees of evil' in their analysis, even though some of them deny this in their views. Through this process of examining and revealing how each author implicitly relies upon the existence of such degrees, we will argue that 'degrees of evil' are an essential and intuitive aspect to any robust account of evil character and is required to articulate many aspects of evil character properly.

Keeping these criteria in mind, we will examine each account of evil character asserted by the five authors. John Kekes argues for a definition of evil character based upon how frequently a person perpetrates evil actions.<sup>1</sup> Colin McGinn asserts an account that views evil character as one's hedonic orientation towards the pain and suffering of others.<sup>2</sup> Daniel Haybron's account of evil character argues that a person is evil if she consistently lacks goodness.<sup>3</sup> Peter Brian Barry argues for a theory of evil character based upon a principle derived from our intuitions, which he calls the mirror thesis, in which the evil person is the opposite of a saintly person.<sup>4</sup> Finally, Luke Russell believes an evil character is someone who has a strong and fixed disposition to perform evil actions when she is freely able to act.<sup>5</sup>

This project will examine and analyze each of these five accounts of evil character against our three criteria to establish whether they are, in fact, robust accounts. Many of these examinations will consist in critiquing each of these accounts in ways that will expose their conceptual errors. Each of these accounts fails to be properly robust in a variety of ways - some critical and insurmountable, some non-critical and recoverable. One criticism that recurs in the critique of these theories is that they cannot account for one or the other of two types of evil character. Daniel Haybron refers to the first type of evil character as the malicious quadriplegic, and it will be used to expand upon his critique of Kekes's frequent evil-doing account.<sup>6</sup> It will also be used to develop a counter-example to Russell's view. Against the other three accounts, a different type of character is needed for a proper critique. This new type of character, the perverse benefactor, will first be argued for, and presented, as a type of evil character that is intuitively acceptable. Afterwards, during our examination chapters, the perverse benefactor will be used to demonstrate how many of our accounts utterly fail to recognize it as an evil character.

This project will follow a specific outline. First, we will present the perverse benefactor character-type in detail. The following five chapters will then examine the authors' accounts of evil character. Finally, for the conclusion, a summary is given of all the arguments made in this project, and a sketch of what a robust account of evil character might look like is provided.



## Notes

---

<sup>1</sup> John Kekes, *Facing Evil*, (West Sussex, Chichester: Princeton University Press, 1990)

<sup>2</sup> McGinn, Colin, *Ethics, Evil and Fiction* (Oxford: Clarendon Press, 1997)

<sup>3</sup> Daniel Haybron, "Consistency of Character and the Character of Evil." in *Earth's Abominations: Philosophical Studies of Evil*, ed. Daniel Haybron. (Amsterdam: Rodopi, 2002)

<sup>4</sup> Peter Brian Barry, "Moral Saints, Moral Monsters, and the Mirror Thesis." *American Philosophical Quarterly* 45 No.2, 2009

<sup>5</sup> Luke Russell, "Dispositional accounts of evil personhood," *Philosophical Studies* 149, No 2, June 2010

<sup>6</sup> Daniel Haybron, "Moral Monsters and Saints," *The Monist* 85, No. 2, 2002

## Chapter 2: The Perverse Benefactor

### 2.1 — A New Example

The purpose of this chapter is to present the perverse benefactor as a type of evil character, although it was not previously considered evil in the accounts of Kekes, McGinn, Haybron, and Barry. In fact, the perverse benefactor is a counter-example to these authors' accounts.

In order to make plausible the perverse benefactor as a type of evil character, we will argue that our common intuition can recognize the perverse benefactor as evil. This argument will be supported through the cultural evidence of a commonly held folk saying. In addition, we will illustrate that the perverse benefactor is a coherent and intuitively acceptable type of evil character. The perverse benefactor does not possess inherent contradictions, nor does it deviate from our intuitions of evil character. In short, we will argue that the perverse benefactor is intuitively and logically plausible. To further explicate this character, we will present an abstract model of the perverse benefactor, as well as more concrete character examples from both fiction and reality.

Finally, we will defend the concept of the perverse benefactor against some criticisms of its intuitive plausibility. It is important to note before moving on that this chapter is meant only to present an example of a specific type of evil character so that it can be used to critique the accounts of Kekes, McGinn, Haybron, and Barry later in this project; it is not meant to provide a robust account of evil character on its own.

## 2.2 —The Perverse Benefactor

Consider the old folk saying, “The road to hell is paved with good intentions.” According to the *American Heritage Dictionary of Idioms*, this proverb was originally coined by Saint Bernard of Clairvaux in 1150 CE.<sup>1</sup> In its original form, Saint Bernard's proverb—“L'enfer est plein de bonnes volontés et désirs”—literally translates into the phrase, “Hell is full of good wishes and desires,” suggesting the existence of an evil character that deserves Hell not because of malevolent thoughts or actions, but rather as a result of good desires.<sup>2</sup> This proverb is integral to our understanding of the perverse benefactor; therefore, to better understand the sort of character this proverb is suggesting, we should examine the elements of the folk saying.

Three components constitute this proverb: Hell, the full nature of Hell, and the contents that fill Hell. In order to understand the first of these components, the concept of Hell must be clarified. Typically, “Hell” is understood in one of at least three ways: as a state of torment, as a place in which torments take place, or, more specifically, as a religious place of torment for which evil people are destined. Given that Bernard of Clairvaux was a twelfth-century saint, one can safely assume that the proverb refers to the biblical concept of Hell. Thus, the first component of the proverb speaks of Hell as a place for evil people: those who have committed heinous acts and are thus deserving of Hell, or, eternal torment.

The second component of the folk saying proposes that Hell is full of something. In the context of Saint Bernard's proverb, this second component refers to the elements with which the biblical Hell is replete, or which characterize most of its populace. Since

Saint Bernard was likely referring to the biblical idea of Hell—a place for evildoers alone—this second component suggests the existence of a characteristic shared by many, or all, of the evildoers in Hell.

Finally, the third component of the proverb reveals that the characteristics that fill Hell are, in fact, good wishes and desires. More specifically, it seems safe to infer from the wording of the folk saying that the ubiquitous features amongst the evildoers in Hell—indeed, the very features that led them to Hell in the first place—are their good wishes and desires.

If this analysis of the old folk saying is correct, the type of character it suggests is an evil character whose evil acts are motivated by good desires. Generally speaking, most people commonly think of evil people as those with evil desires who intend to do evil things and act accordingly; in other words, they think of villains or moral monsters who desire only to rob, injure, kill, and destroy. Yet, this proverb suggests a person who desires good things and yet is still evil and deserving of a biblical Hell. As this type of person desires what is good but does evil, we will henceforth refer to it as the perverse benefactor.

The perverse benefactor presents a complicated notion of evil character, defying the simplicity of the assumption that an evil character is always monstrous. In other words, the perverse benefactor's very existence challenges the idea that an evil person must desire what is evil and do evil to achieve it. It similarly complicates the notion that moral saints desire what is good and perform good acts to achieve their desires. Thus, the perverse benefactor highlights a more complicated relationship between what she desires

and what she does to achieve her desires.

As a brief example of the perverse benefactor, consider the tragic cases of some mothers suffering from postpartum depression. In certain cases, some mothers believed that the world was too harsh a place in which to raise their children, and so, to save their children from a lifetime of hardship and suffering, they have desired to kill their own children. These mothers desired what was best for their children—broadly speaking, that their children should have few hardships and little suffering—but they believed that the best way to achieve this goal is to kill their children. It is specifically this sort of perverse and wicked thinking that the perverse benefactor type of character is meant to embody.

However, not all perverse benefactors are mothers suffering from postpartum depression. Therefore, the purpose of this chapter is to present and argue for a broader character type that models the foundational perverse relationship between a person's good desires and her decision to fulfill that desire through evil actions. What follows is an abstract model of the perverse benefactor character type, which will be followed by a specific example of one such person in the real world.

### **2.3 — The Characteristics of a Perverse Benefactor**

Extrapolating from the above folk saying, one can infer that a perverse benefactor must possess an effective desire to promote the wellbeing of another individual or group. By effective desire, I mean a desire that principally moves one to action.<sup>3</sup> This effective desire should not include the promotion of one's own wellbeing; at least, one's own wellbeing should not be what makes the desire an effective one. If someone behaves

altruistically only when she can also benefit herself, it is unlikely that the goal of the other is her principal motivation. The mothers in the example above certainly were not motivated by desires for their own wellbeing; they were moved by a desire for their children's wellbeing. Accordingly, to be a perverse benefactor, one must be moved by such a desire to commit evil actions. Further, it appears a perverse benefactor must genuinely believe that her intended evil actions are the right thing to do, or at least are morally permissible, in service of her desire(s). The depressed mother would not be an example of a perverse benefactor if she did not want to promote her child's wellbeing by killing it, and did not believe that killing it was a morally permissible thing to do. In order to understand the perverse benefactor as a type of evil character, it is essential that one hold a definition of evil action that allows for the necessary perversion to develop that leads to the fulfillment of altruistic desires through evil actions. In other words, the perverse benefactor relies upon the understanding that evil action contains a harm component. Under that sort of definition, the perverse benefactor establishes the existence of a perverse connection between altruistic desires and evil actions—that is, actions that cause harm. What precisely constitutes an evil action still remains a debated issue, and therefore the model of the perverse benefactor will simply have to rely upon whatever account of evil action is determined. Because the provision of a proper account of evil action is beyond the scope of this project, an understanding of the essential requirement for a perverse relationship between altruism and action will have to suffice.

The broad model of the perverse benefactor depicts a character that desires to promote the wellbeing of others and yet fulfils those desires through evil actions. In

essence, the relationship between this character's altruistic desires and her intentions to act wickedly to fulfill those desires is a perverse one. This model is easy to understand, and it allows one to recognize perverse benefactors readily operating in the real world. Yet, one thing remains unclear from this model, and that is the question of how any person could believe that performing evil actions would fulfill her altruistic desires.

When attempting to reconcile the perverse relationship that exists between a perverse benefactor's desires and actions, one must consider the several possibilities that could provide a satisfactory explanation. On one hand, perverse benefactors could rationalize their actions according to the old adage, the ends justify the means. They could therefore believe that their altruistic desires morally justify their evil actions, perhaps thinking that evil actions might be more efficient when fulfilling their desires. On the other hand, perhaps they are simply ignorant of the fact that their actions are actually evil. They might believe that there is nothing morally wrong within their planned courses of action.

Each of these explanations relies upon the common assumption that the perverse benefactor does not—or will not—confront the evil nature of her actions. In both of these potential explanations, the perverse benefactor believes that her actions are morally justifiable, either because she possesses a poor estimation of moral quality or she is ignorant thereof. David Jones, for one, suggests that self-deception may be the best explanation behind the moral justification of one's evil actions.<sup>4</sup> Jones reflects upon the work of Mike Martin, who explains the concept of self-deception and how it could present itself.<sup>5</sup> Martin views self-deception as the refusal to confront facts squarely or

straightforwardly, as they pertain to some particular circumstance.<sup>6</sup> Martin's work pertains to the perverse benefactor character type insofar as perverse benefactors could be deceiving themselves into thinking that their actions are morally justified, by refusing to confront the evil nature of their actions straightforwardly. The concept of self-deception is present in a variety of rationalizations that could be used to morally justify one's actions to oneself. For instance, one perverse benefactor could believe that the ends justify the means because she has deceived herself into thinking that this is the case. Meanwhile, another perverse benefactor could be ignorant of the evil nature of her actions because she has deluded herself into thinking that they are good and wholesome actions, given her particular circumstances. Under any of these, or similar, rationalizations the self-deceptive person lacks adequate justificatory reasons for believing her evil actions are morally right or permissible, thereby necessitating such rationalizations in the first place. Martin and Jones argue that it is the self-deceptive person's lack of adequate justificatory reasons—i.e., good reasons—for believing her actions morally justified that makes her morally blameworthy for thinking as she does.<sup>7</sup>

Given Martin's concept of self-deception, we can reasonably describe the perverse benefactor as a type of evil person who is effectively motivated by her desire to promote the wellbeing of others through evil actions that she has deceived herself into thinking are morally justified. For example, the mother suffering from postpartum depression who believes that the best thing she can do for her children is to kill them has clearly been deluded into this belief by her depressed state.

Before moving forward, it would be beneficial to consider an example of a



historical figure that is also a paradigm example of a perverse benefactor. David Jones writes of Adolf Hitler as an evil person who is self-deceived, and as we go into detail discussing Hitler, the reader should consider how closely he matches the proposed model of a perverse benefactor.<sup>8</sup>

Hitler is intuitively considered to be an evil character: that much is generally not disputed. Moreover, Hitler is frequently depicted as a moral monster moved by evil desires to commit acts of genocide toward people of various ethnicities. David Jones argues that Hitler's hatred for the Jewish people, among others, is the result of a life-long campaign of self-deception, and that Hitler is culpable for his self-deception.<sup>9</sup> Hitler's self-deception essentially led him to believe that the Jewish people were a threat to Germany, Europe and possibly even the world.<sup>10</sup> As Jones argues, this belief motivated Hitler to help the German people overcome this perceived danger.<sup>11</sup> I contend that, when combined with the self-delusion that European Jews were the root cause of Germany's national woes, Hitler's desire to help the German people languishing from various penalties from the First World War served to create a perversion. His otherwise good desires led him to attempt an evil action: the elimination of millions of Jews and people of various other ethnicities, the Holocaust. Hitler's actions themselves are not the focus of this example, though; rather, we must consider what motivated him to commit them. Jones argues that Hitler was directed, in large part, by his own self-deluded beliefs.<sup>12</sup> I believe that the perverse relationship between Hitler's good effective desire to help others and his attempts to do so through evil actions—in this case attempted genocide—makes him a concrete example of a perverse benefactor. Jones argues that Hitler is culpable for

his self-deception, making him morally responsible for his false beliefs.<sup>13</sup> If true, this classifies Hitler as a sort of evil person motivated by good desires instead of evil or malicious ones. At the very least, Hitler is an evil person best described not as a moral monster—an evil character whose evil actions are motivated by evil desires—but as a perverse benefactor.

One can easily imagine Hitler as a moral monster. However, such a reckoning of his evil character would involve a misinformed analysis of his motivations. Hitler's hatred of the Jews could be viewed as nothing more than a mere brute fact, and his wish to destroy them nothing more than the result of that hatred. Hypothetically, Hitler could also have been a person of a very different character. For instance, if he saw the German people languishing under the penalties of the First World War but did not blame European Jewry for their problems. In either situation, Hitler would not demonstrate the perverse relationship extant between good effective desires and evil actions that he demonstrated so effectively in our real history. Therefore, Hitler is an evil person best described as a perverse benefactor; likewise, since Hitler is intuitively acceptable as an evil person, the model of the perverse benefactor is also clearly intuitively plausible.

Now that we have described the character of the perverse benefactor and have also considered an example of a perverse benefactor who is undeniably intuitively discernable as evil, we will turn to some possible criticisms of the perverse benefactor model and will offer rebuttals to them.

## **2.4 — Criticism against Good Desires**

A critic could suggest that the perverse benefactor is not actually evil and could make the argument that only non-evil characters have good effective desires and therefore evil people cannot have them. This concern is understandable, since a robust account of evil character ought to determine that only the very worst sorts of people can achieve the sufficient conditions for evil to obtain. According to this criticism, the perverse benefactor example, if true, would suggest that the sufficient conditions for evil are too low, allowing some merely very bad characters to be considered evil when they, in fact, are not.

The above criticism assumes the existence of a maxim: that evil people do not have good desires. Yet, no such maxim actually exists. Evil people are commonly conceived of as cackling, maniacal, moral monsters that only desire evil things such as pain and suffering. However, evil characters can be seen as having non-evil effective desires, a subset of which might be good effective desires. Furthermore, there is nothing logically inconsistent with evil persons being moved by good desires. For instance, it is not impossible to believe that an evil person could easily and intuitively desire the wellbeing of her own children. A moral monster could even, in a more minimalist sense, truly desire good for her pet. If evil persons are not permitted to have any good effective desires, we run the risk of setting the sufficient conditions for evil too high. That would be problematic, as it would contradict our intuitions that at least some evil people exist. Yet, since we know that evil people are at least occasionally moved by good desires, in order to make the perverse benefactor intuitively unacceptable, we would need to be willing to accept the maxim that evil people never have good effective desires. Since

there is no good reason to believe that evil people are not moved by at least some good desires some of the time, there also is no good reason to think that there could not be an evil person who has effective good desires the majority of the time. If we allow evil persons even some good effective desires, then the perverse benefactor stands as an example of a logically possible type of evil character.

## 2.5 — Summary

The reader should keep in mind that the perverse benefactor is intended to be an example with which the accounts of John Kekes, Colin McGinn, Daniel Haybron and Peter Brian Barry will be critiqued. The perverse benefactor is at least acceptable enough to our common intuitions to serve in that capacity. The perverse benefactor is also coherent with our intuitions about other types of evil characters. Additionally, the model of the perverse benefactor provides a more adequate explanation of specific types of evil characters, such as Hitler or infanticidal mothers, than is provided by accounts based solely on models of moral monsters, for instance. For the purposes of this project, the perverse benefactor will be used as a model to illustrate that evil characters that do not desire evil for its own sake are entirely unaccounted for in these four authors' accounts. Beginning in the following chapter, the five accounts of evil character proposed by the aforementioned authors will be examined.

## Notes

---

<sup>1</sup> Christine Ammer, *The American Heritage Dictionary of Idioms* (Houghton Mifflin, 1997)

---

<sup>2</sup> By good desires I mean desires for the good of others or altruistic desires, not desires that are free of corrupting external influences.

<sup>3</sup> Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy* Vol 68 No 1 (Jan 1971), 8

<sup>4</sup> David H. Jones, *Moral Responsibility in the Holocaust: a study of in the ethics of character*, (Rowman & Littlefield. Oxford, Uk 1999), 79-95 & 121-142

<sup>5</sup> Mike W. Martin, *Self-Deception and Morality* (Lawrence: University Press of Kansas, 1986), 13

<sup>6</sup> Martin, *Self-Deception*, 13

<sup>7</sup> David H. Jones, *Moral Responsibility in the Holocaust: a study of in the ethics of character*, (Rowman & Littlefield. Oxford, Uk 1999), 85-90

<sup>8</sup> Jones, *Moral Responsibility*, 137-142

<sup>9</sup> *Ibid.*, 121-142

<sup>10</sup> *Ibid.*, 129-130 and throughout

<sup>11</sup> *Ibid.*, 130

<sup>12</sup> *Ibid.*, 130

<sup>13</sup> *Ibid.*, 137-142

## Chapter 3: John Kekes's Frequent Evil-Doing Account

### 3.1 — The Account

This chapter will examine John Kekes's frequent evil-doing account, in which Kekes argues that the overabundance of evil present in the world is disproportionate to the number of people classified as evil.<sup>1</sup> In an attempt to explain this disparity, he argues that evil actions are not performed solely by a relatively small number of moral monsters.<sup>2</sup> Instead, Kekes believes that defining only the most morally corrupt characters as evil is a mistake, as it excuses innumerable other evildoers from being considered evil.<sup>3</sup> Furthermore, Kekes believes such a discrepancy exists because people are uncomfortable believing that the world contains a vast number of evil characters, and so we commonly excuse the evil actions of people who do not quite fall into the same category as moral monsters.<sup>4</sup> Because there are some who would disagree with this reasoning, Kekes has delineated two different approaches to evil character: the "hard reaction" and the "soft reaction." The "hard reaction" refers to the assumption that the large amount of evil present in the world is a result of the many evil characters also present.<sup>5</sup> Alternately, the "soft reaction," which is a popular concept among some philosophers, pertains to the belief that "evil" is a term that should be reserved for only the morally worst characters.<sup>6</sup>

Kekes himself embraces the hard reaction to evil, as he believes that the abundance of evil present in this world is strong evidence for the existence of a great number of evil characters.<sup>7</sup> Furthermore, he is able to uphold the hard reaction by providing a definition of evil action as that which causes sufficiently severe and

inexcusable undeserved harm, actions that he believes happen all the time.<sup>8</sup> Kekes's original account of evil character asserted that a person could be considered an evildoer if she was a frequent cause of inexcusable severe undeserved harm.<sup>9</sup> However, he later adjusted his account to posit that an evildoer is an individual who is malevolently motivated to cause such harm and does so frequently.<sup>10</sup> Kekes then identifies frequent evildoers as evil characters, correspondingly allowing his account to be termed a frequent evil-doing account.<sup>11</sup>

In order to comprehend properly Kekes's understanding of evil character, one must consider his definition of harm. According to Kekes, harm can be any painful or unpleasant sensation or experience—either mental or physical.<sup>12</sup> The harm caused, though, must be severe enough to create distress, rather than mere frustration or annoyance. In addition, it must be undeserved; just punishments or appropriate consequences do not constitute an acceptable type of harm. Finally, the harm must be morally inexcusable. In other words, the action that brings harm about cannot be morally justified. For example, the execution of a convicted murderer for his crime would not be considered the type of harm that would classify as the performance of an evil action. On the other hand, the cold-blooded murder of an innocent man certainly would be considered evil.

Kekes's account argues that the status of a character as evil depends upon the frequency at which she commits harmful acts.<sup>13</sup> While malevolent motivation may determine that a harmful action is in fact an evil action, we cannot determine the condition of one's character unless we know how frequently the action is committed.<sup>14</sup>

For instance, in Kekes's view, a malicious quadriplegic who would gleefully murder everyone but is physically incapable of doing so would not be considered an evil character because she cannot cause harm at all, let alone frequently.

Finally, Kekes also endorses the concept of degrees of evil character, and he claims that his account accommodates this concept by allowing for increases and decreases in the frequency, severity and malice of evildoers' actions. In essence, he argues that those who cause harm more maliciously, severely, or frequently than others are evil to a greater degree than those others.<sup>15</sup> According to Kekes's view, then, moral monsters are evil characters who commit greater degrees of harm more frequently than the majority of other evil people; yet, they are not the *only* evil people in existence, despite the assertions of some, "soft reaction," philosophers.

As this project is primarily concerned with finding a robust account of evil character, it is necessary that we highlight what Kekes's account describes as the necessary and sufficient conditions for evil character before we can critically examine his view. Kekes's account is that being a frequent source of malevolently caused significant undeserved inexcusable harm is sufficient and necessary for evil character. If an individual is not a frequent source of such harm, that person is not evil, in Kekes's view. The following section will examine Kekes's account to see if it is a robust one with intuitively plausible conditions for evil character. Two of the criticisms that will be used to examine Kekes's account come from Daniel Haybron, one of the authors whose account will be examined later. The last criticism arises from Kekes himself; more specifically, it is prompted by his inclusion of a malevolent motivation component in a



later adjustment to his account.

### **3.2 —How Frequently is Frequently Enough**

In his first criticism of Kekes's account, Daniel Haybron asks how frequently one must cause undeserved harm in order to be evil.<sup>16</sup> Haybron's question highlights the vague nature of Kekes's definition of evil character, since Kekes does not delineate exactly how frequently one must cause such harms if one is to be considered evil. While Kekes's use of the term frequent suggests that an evil character must commit more than one malevolent harmful act, the minimum amount of harmful actions sufficient for evil character to obtain remains open to interpretation. The amount must be greater than one, though, because, if one were to argue that a single transgression is sufficient for evil character to obtain, then the proposed account would not be a frequent evil-doing account.<sup>17</sup> However, the problem with Kekes's view is that one cannot reasonably provide a plausible number of transgressions that would determine the evil nature of a character. For instance, it is completely unjustifiable to say that twenty evil actions per month qualifies one as evil, as opposed any other possible frequency. A proponent of Kekes's view ought not to be comfortable with such a vague and implausible basis for defining evil character.

### **3.3 — The Malicious Quadriplegic**

For his second criticism, Daniel Haybron utilizes an example of a malicious quadriplegic who desires to watch and relish the pain and suffering of others.<sup>18</sup> Haybron's

example prompts questions about how one's physical abilities effect one's moral character, such as whether the inability to cause undeserved harm makes a malicious quadriplegic a more moral person. Even though the quadriplegic would cause harm if she were able, the fact that she is incapable of doing so prevents her from being considered an evil character on Kekes's account.

By presenting a malicious quadriplegic as an intuitively evil character, Haybron proposes that being the frequent cause of undeserved harm is not, in fact, an intuitively plausible necessary condition for evil. If the malicious quadriplegic is acceptable as evil to our intuitions, Haybron's example shows that an account of evil character must not include conditions for evil that are contingent solely upon one's actions.

### **3.4 — The Specific Motivation Criticism**

When Kekes adjusted his account in order to allow that evil actions—that is, actions that cause significant inexcusable undeserved harm—are only evil when fueled by malicious motivations, he provided the grounds necessary for one final criticism of his account. Above all, the adjustment to Kekes's account prevents the perverse benefactor from being considered an evil character. The perverse benefactor, who was discussed at length in the previous chapter, is an intuitively plausible evil character who is effectively moved by benevolence, rather than malevolence. The failure to accommodate the perverse benefactor as evil further demonstrates that Kekes's account is not robust, since any robust accounting of evil character should be able to accommodate all types of evil characters.

A proponent of Kekes's account could suggest that the perverse benefactor is not an evil character because she does not perform evil actions, since, in his view, evil actions can only be performed from malevolent motivations.<sup>19</sup> However, this manner of rebuttal was addressed in the previous chapter, in which it was argued that there is no good reason to think that altruistic motivations could not move one to act evilly, or, to use Kekes's terminology, to cause undeserved harm. Therefore, since the perverse benefactor represents one type of evil character that Kekes's account simply does not accommodate, his account is not a robust view of evil character

### **3.5 — Summary**

In this chapter, we examined the frequent evil-doing account developed by John Kekes, which defines an evil person as one who malevolently causes severe and inexcusable undeserved harm frequently. Kekes's account was chosen to exemplify an account of evil character that strictly requires an external action component in its analysis. Criticisms of Kekes's account demonstrate its inability to meet the criteria necessary for the formation of a robust account of evil character. For instance, the account lacks plausible sufficient and necessary conditions for evil character, by making its only standard—that of frequency—vague and unintuitive. In addition, Kekes's account can accommodate only evil characters that are malevolently motivated, whereas a robust account ought to be able to accommodate all evil characters, such as those same ones that are altruistically motivated.

By demonstrating the problems inherent in Kekes's account, this chapter has

illuminated ways in which a robust account of evil character should be developed. First, a robust account should not require the performance of evil actions. Rather than assuming that one's actions causally determine one's character, one should acknowledge the opposite—that an individual chooses to behave in certain ways because one's character inclines her to do so. Character, rather than action, should function as the determinant. In the words of Luke Russell, a robust account of evil character must make sense of the phrase, “he did it because he was evil.”<sup>20</sup> Second, a robust account of evil character must avoid vague components. Philosophers of evil, and indeed everyone, should strive for precision and clarity when attempting to construct meaningful accounts. Finally, greater attention is needed when considering components that enhance the precision of an account. Kekes added a component of malevolent motivation to his account in response to criticisms that his account lacked a psychological component. However, he did not fully consider the possibility that adding this component would inadvertently eliminate at least one possible evil character from his account. This exclusion alone is reason enough to argue that this account is not a robust one.

## Notes

---

<sup>1</sup> John Kekes, *Facing Evil*, (West Sussex, Chichester: Princeton University Press, 1990), 6

<sup>2</sup> John Kekes, *Facing Evil*, 6

<sup>3</sup> *Ibid.*, 6

<sup>4</sup> *Ibid.*, 5-6

<sup>5</sup> *Ibid.*, 6-7

<sup>6</sup> *Ibid.*, 5-6

<sup>7</sup> *Ibid.*, 6-8 Kekes conflates evildoers with evil people in his more recent book. He says: “Evil actions violate these basic limits, and evil people do [evil actions] regularly, excessively, malevolently, and inexcusably.” (Kekes, *Roots*, p.130)

<sup>8</sup> *Ibid.*, 3-7

<sup>9</sup> *Ibid.*, 7

<sup>10</sup> John Kekes, *The Roots of Evil*, (Ithaca NY: Cornell University Press, 2005), 2

---

<sup>11</sup> John Kekes, *Facing Evil*, (West Sussex, Chichester: Princeton University Press, 1990), 7 & John Kekes, *The Roots of Evil*, (Ithaca NY: Cornell University Press, 2005), 2

<sup>12</sup> John Kekes, *The Roots of Evil*, 1

<sup>13</sup> *Ibid.*, p.2 ; John Kekes, *Facing Evil*, 7

<sup>14</sup> John Kekes, *The Roots of Evil*, 2

<sup>15</sup> *Ibid.*, 2

<sup>16</sup> Daniel Haybron, "Moral Monsters and Saints," *The Monist* 85, No. 2, 2002, 264 Haybron criticizes Kekes's original account here. However, the criticism works equally well against Kekes's latest account.

<sup>17</sup> The sort of problem that Kekes is having in defining a specific frequency is like the paradox of the heap.

<sup>18</sup> Daniel Haybron, *The Monist*, 264

<sup>19</sup> John Kekes. *The Roots of Evil*, (Ithaca NY: Cornell University Press, 2005), 2

<sup>20</sup> Luke Russell, "Dispositional accounts of evil personhood," *Philosophical Studies* 149, No 2, June 2010, 232

## **Towards a Robust Theory of Evil Character: Chapter 4**

### **4.1 — Colin McGinn's Affective Account**

Colin McGinn provides an account of evil character based on one's affective response toward pain and pleasure.<sup>1</sup> Two central elements comprise McGinn's view of evil character, the first of which is called the hedonic disposition. A hedonic disposition refers to one's emotional psychology, specifically regarding the manner in which the pain and suffering, or pleasure, of another affects one. McGinn says that morally decent characters feel pleasure when another feels pleasure.<sup>2</sup> In other words, morally decent people empathize with others: they are happy for others when others are happy. Correspondingly, morally decent people are pained when others are pained. McGinn contends that evil characters do not empathize with other people; they experience pain when another feels pleasure, and pleasure when another is in pain.<sup>3</sup> Thus, the first element of McGinn's view to understand is that evil characters have completely reversed hedonic dispositions that cause them to be pained by another's pleasure and pleased by another's pain.<sup>4</sup>

Another element of McGinn's account is his choice to exclude a motivational component from his theory.<sup>5</sup> His exclusion of a motivational component allows for the possibility that some evil people may wish only to observe the pain and suffering of others, rather than to perform actions that cause pain.<sup>6</sup>

Furthermore, evil people, in McGinn's view, do not need to be killers who seek to end the lives of others. In McGinn's view, the experience of pleasure is more important to

an evil person than the ending of a life.<sup>7</sup> Since the pain one can feel ends at death, an evil character does not necessarily benefit from killing others, since her own pleasure would be limited by the death of her victim.<sup>8</sup> Therefore, to be considered an evil character by this account does not imply that one must actively seek to hurt others. Instead, what is of import is that one takes pleasure in the pain and suffering of others, and is similarly pained by the pleasure of others.

To summarize McGinn's account, it is necessary and sufficient for evil character to have a reversed hedonic disposition; in other words, an evil character must experience pain at the pleasure of others and experience pleasure at the pain of others. It is not necessary for evil characters actively to hurt others; merely having a reversed disposition is sufficient.

#### **4.2 — Excluding Motivation**

Daniel Haybron makes this first criticism of McGinn's account. Haybron asserts that the affective account's exclusion of matters of motivation is detrimental to its analysis.<sup>9</sup> He argues that it is important for any account of evil character to explain why that character behaves the way she does. A proponent of the affective account could have suggested that the evil character is motivated by the pursuit of pleasure if McGinn had not specifically excluded matters of motivation from his view. This leaves the affective account in an awkward position, in which it is conceptually possible for an evil person who is greatly pleased by the pain of others to lack the motivation to seek it out. However, the idea that someone—like a sadist—whose greatest pleasure comes from the

suffering of others would never be moved to seek out that pleasure seems intuitively implausible. The implication that this criticism demonstrates is that the affective account lacks a plausible formulation of evil character, since a robust account should be able to explain why evil people behave the way they do.

With this criticism in mind, it could also be argued that without a consideration of matters of motivation, the mentally ill might qualify as evil according to McGinn's account. It seems that in some way an evil person must choose to behave immorally. Without a motivation component in an analysis of evil character, those who cannot help themselves, such as psychotic patients, may be considered evil. For instance, it is possible that someone suffering from some form psychosis might—for that reason—take pleasure in the suffering of others. This psychosis would seem to exculpate the perpetrator from being considered evil; however, that would not be the case, given the current structure of the affective account.

### **4.3 – Frequency and Intensity**

The next criticism also comes from Haybron. He suggests that the affective account is incapable of offering a firm distinction between very bad and truly evil characters. Haybron's point attempts to illustrate the difficulty of establishing just how frequently one must experience pleasure at the pain of others, and vice versa, to be evil. In a manner similar to a criticism of Kekes's theory, this question highlights the difficulty the affective account faces when establishing a clear sufficient condition for evil character. If an evil person must always have the affective responses McGinn suggests,



the bar might be set too high. Even evil people such as Adolf Hitler experience normal hedonic reactions. Haybron's criticism concerning the frequency with which one must have a reversed hedonic experience can be paralleled by asking how much pleasure one must experience at another's pain to be evil. It is an acceptably intuitive fact that certain hedonic experiences are more or less pleasurable than others, and McGinn's account is equally vague about how strongly one must be pleased by another's pain to be evil. After all, even morally decent people take a little pleasure in the pain of others some of the time, such as in cases of *schadenfreude*. Therefore, it appears that McGinn's account needs to be clearer about what constitutes a sufficient hedonic disposition if one is to be evil. A robust account of evil character ought to have clear and intuitively plausible conditions, and McGinn's view seems to fall victim to problems similar to those that made Kekes's account implausible.

#### **4.4 — Evil without a Reversed Disposition**

Peter Brian Barry makes our final criticism. Barry argues that the entire affective account could be refuted with a single example of an evil person who does not have a reversed hedonic disposition, and to prove this point he suggests a thought experiment concerning Dennis Rader, the BTK Killer.<sup>10</sup> Barry asks us to suppose that Rader may have become jaded after so many killings, no longer taking any pleasure in them but continuing to murder people out of habit. Barry suggests that Rader's lack of derived pleasure would scarcely make him a better person. For another example, recall the perverse benefactor discussed in the second chapter. The affective account simply cannot

accommodate a person who lacks a reversed hedonic disposition, such as Barry's hypothetical Dennis Rader, let alone an evil person who possesses a normal hedonic disposition. The perverse benefactor's desire for good clearly demonstrates that it has a normal affective nature. If the perverse benefactor is an intuitive example of a type of evil character, then, according to Barry, the affective account is refuted entirely.

#### **4.5 — Summary**

This chapter has examined Colin McGinn's affective account, according to which evil people are those who have hedonic dispositions that are reversed compared to those of morally decent people. Simply put, the evil person finds pleasure in the pain of others and pain in the pleasure of others. McGinn further implies that there are no degrees of evil character, and he denies that motivation is an important factor to consider when defining evil character.

McGinn's account cannot establish an intuitively plausible set of conditions for evil character; he does not allow for a firm distinction between the occasionally bad and the truly evil. Similarly, his account may not be able to distinguish between those people who are not morally responsible for their evil actions and those who are. Also, McGinn's account cannot intelligibly determine what level of reversed affective response is sufficient for evil, or how frequently one must experience such a reversed affective response. Finally, this account is quite possibly refuted by Barry's criticism, that evil might be done affectlessly, which is supported in part by the example of the perverse benefactor.

## Notes

---

<sup>1</sup> McGinn, Colin, *Ethics, Evil and Fiction* (Oxford: Clarendon Press, 1997)

<sup>2</sup> McGinn, Colin, *Ethics*, 61-63

<sup>3</sup> *Ibid.*, 61 & 65-59

<sup>4</sup> *Ibid.*, 62

<sup>5</sup> *Ibid.*, 66

<sup>6</sup> *Ibid.*, 66

<sup>7</sup> *Ibid.*, 68

<sup>8</sup> *Ibid.*, 68

<sup>9</sup> Daniel Haybron, "Moral Monsters and Saints," *The Monist* 85, No. 2, 2002, 268

<sup>10</sup> Peter Brian Barry, "Moral Saints, Moral Monsters, and the Mirror Thesis," *American Philosophical Quarterly* 45 No.2, 2009, 166

## Chapter 5: Daniel Haybron's Consistency Account

### 5.1 — The Account

Daniel Haybron developed his consistency account to illustrate how definitions of evil character that require a specific trait, or set of traits, might be fundamentally erroneous, and he convincingly notes that accounts dependent upon such traits or sets of traits are frequently too narrow.<sup>1</sup> Inevitably, counter-examples emerge to prove that trait-based accounts are incomplete. Thus, Haybron suggests that evil character should not be defined by a specific characteristic, or even a select set of characteristics, but that an evil person is someone who is consistently and entirely evil.<sup>2</sup> Thus, Haybron puts forward an account that evaluates one's entire character in an attempt to determine whether or not it is consistently malignant, rather than setting some limited standard of sufficient extremity.

According to the consistency account, a person is evil if she lacks a significant good side.<sup>3</sup> To illustrate this claim, Haybron discusses the fictional character Chad, from the film *In the Company of Men* (1997), and asserts that Chad is an intuitively evil character.<sup>4</sup> However, Haybron does not see Chad as extremely evil, and he describes the man as "cruel, but not Satanically so."<sup>5</sup> Thus, in an attempt to articulate what makes Chad evil, since he is not extremely cruel, Haybron argues that Chad's evil character is found in his thorough and consistent lack of a good side.<sup>6</sup>

Haybron elaborates on his account's condition for evil character by presenting some different ways the lack of a good side could be phrased or considered.<sup>7</sup> The evil person—in this case Chad—lacks a good side because he is not positively moved by the

good. By this, Haybron means that the evil character's affective response towards the good is not a positive one. Simply put, morally good ideals do not affect him. By the good, Haybron means something akin to a Platonic sense of the good, or a moral ideal. Alternately, the good can also be considered in terms of virtue and vice. Haybron explains that if a person's virtues and vices are taken to be signs of one's orientation with the good, an evil person is one that thoroughly and consistently lacks virtue. Another way of phrasing Haybron's claim is by saying that evil people are not attracted to the good. The point Haybron means to assert is that evil people are not moved or motivated to pursue the moral ideal. Essentially, evil people are thoroughly and consistently unmoved and unmotivated by the good; they simply do not care for it.<sup>8</sup>

The extent to which evil people must be unmoved by the good is a question Haybron debates.<sup>9</sup> He presents two possible views regarding the manner in which the evil person is unaligned with the good, which he terms a weaker view and a stronger view. According to the weaker view, the evil person needs only to be mostly unaligned with the good. In this view, evil people have significant leeway in the extent to which they can be moved by the good. Accordingly, an evil person could maintain some minor moral motivations, such as caring—or only caring insignificantly—for one's pet animals or close family. The stronger view demands that evil characters are wholly or entirely unmoved by the good. On the stronger view, the evil character has no moral compunctions whatsoever, and any appearances of morality she might have surely result from amoral motivations. Haybron prefers and endorses the stronger view.<sup>10</sup>

Haybron's goal is to establish a firm distinction between mere ne'er-do-wells and

the truly evil.<sup>11</sup> Thus, he hinges his analysis upon the evil person's consistent lack of motivation by the good. This permits non-evil people to behave immorally on occasion, even though they lack the consistency of character that the truly evil allegedly possess. This view also permits evil people to behave morally on occasion or even very frequently, despite the strong view's assertion that evil people do not react to moral motivations. Even on the strong view, the evil person is permitted to slip up, as it were, and experience an abnormal moment of moral motivation as long as she remains consistently unmoved in such situations. Additionally, the evil person could appear to act morally, being moved instead by amoral concerns about aesthetics or prudence instead of moral ones. By allowing evil people to avoid specific trait criticisms, Haybron makes the consistency account more intuitively plausible and he believes this allows his account to avoid the problems faced by trait-based accounts.<sup>12</sup>

To summarize Haybron's account, it is necessary and sufficient for evil character when one's character is thoroughly and consistently unmoved—positively—by the good or the moral ideal.<sup>13</sup> Haybron's account alleges a firm distinction between the very bad and the truly evil by asserting that only truly evil characters thoroughly and consistently lack a positive motivation from the good. Haybron obviously permits evil characters in his analysis to be motivated negatively by the good to behave immorally. However, it should also be noted that Haybron strongly rejects the idea that there could be degrees of evil character in his, or any, account. He claims that labeling someone as “very very evil,” as opposed to just “evil” is meaningless.<sup>14</sup> According to the consistency account, a person who is thoroughly and consistently unmoved by the good cannot be further unaligned

with the good, and therefore degrees-of-evil terminology cannot be used to describe certain people as more evil than others.

## **5.2 — An Unclear Condition**

On Haybron's account evil people are wholly and consistently unmoved by the good. But what he means by consistently remains vague. To be consistent, an evil person may need to be wholly unmoved by every particular good, so as to be wholly uncaring about anyone's well-being or upholding any specific virtue. Let us call being unmoved by every particular good perfect consistency. However, if he means that to have a consistent character one must be unmoved by every particular good, then his account will violate our intuitions that some evil people exist.<sup>15</sup> It seems highly unlikely, to the point of absurdity, that any human person could be unmoved by every possible moral good. Alternately, if his account does not require perfect consistency for evil character then we must ask how often must we be unmoved by the good, to be evil. It is hard to see how he could adequately specify the number of particular goods we would need to be unmoved by in order to be evil.<sup>16</sup> Thus, Haybron's notion of consistency is woefully unclear and requires a more detailed definition.

## **5.3 — Lacking descriptive power**

Peter Brian Barry offers a criticism of Haybron's account in which he notes that it "says far too little about what evil persons *are* like aside from pointing out what they

*lack.*<sup>17</sup> To illustrate this point, Barry discusses Rousseau's noble savages, who lack moral relationships and a knowledge of good and evil. Barry points out that noble savages are neither good nor evil, since they lack moral knowledge. This presents a problem for Haybron's account, as Haybron stipulates that a lack of moral motivations is necessary for evil to obtain.

Barry argues that our intuitions require evil people to be motivated in a particular way; we would like to believe that they do not merely avoid doing what is right, but that they actively seek to do what is evil and wicked. Barry's criticism demonstrates that the consistency account is unable to explain the source of this intuition. In short, Haybron appears to have confused amoral actions with immoral behavior, or being amoral with being immoral.

#### **5.4 — No Degrees-of-evil**

Haybron's account should further be criticized for its denial of the existence of degrees of evil.<sup>18</sup> As illustrated by previous chapters, degrees of evil are an integral and intuitive element in any robust account of evil character. Yet, while describing the foundations of his account, Haybron specifically denies the need for degrees of evil within any account of evil character. If true, Haybron's claim renders his account unacceptable.

For instance, when Haybron provides Chad as an example of a person of evil character, he notes that one of Chad's defining features is his lack of extreme cruelty.<sup>19</sup> In Haybron's view, Chad is not evil because he is extremely wicked, but because he is



consistently wicked.<sup>20</sup> As noted previously, Haybron describes Chad as “cruel, but not Satanically so.” Haybron’s claim that Chad is less evil than a Satanically cruel figure, which is more evil, implies the existence of varying degrees of evil. While Haybron intended his comparison between Chad and Satan to illustrate how irrelevant the magnitude of cruelty is when determining the evil nature of a character, his comparison actually suggests the existence of a common intuitive understanding that some people are more evil than others.<sup>21</sup> As his example ironically demonstrates, greater and lesser evil are intuitively intelligible terms.

Despite Haybron’s claims to the contrary, his account needs to consider degrees of evil. Haybron states, “to be evil is to be disposed to be neither moved nor motivated (positively) by the good to a morally significant extent.”<sup>22</sup> A person’s disposition and the extent to which it is aligned with something are both variable elements, each having degrees of magnitude. How disposed an individual is toward one thing over another can change, or vary, quite easily. Equally variable is the extent to which one is moved by moral stimuli. In essence, people have the capacity to care about different things to varying degrees. However, if an account is contingent upon one’s affective responses, as Haybron’s is, that account must explain the extant differences between those who care a little and those who care a lot.<sup>23</sup> Haybron’s clarification regarding evil disposition does little to remove the implied degrees-of-evil terminology from his account.

Haybron’s analysis acknowledges the existence of at least two different kinds of evil persons. The first kind is similar to Chad: she is unmoved by and uncaring for the good. The second type is moved negatively by the good and seeks to oppose it. In regards

to his comparative example, Haybron claims that this second kind of character is likely to be recognized as more “like Satan” than like Chad, who, as Haybron argues, is not as wicked as a Satan figure.<sup>24</sup> Without the inclusion of degrees-of-evil terminology, the statement that Chad is not as wicked as Satan is incoherent. The distinction between greater, or lesser, types of evil characters is implausible if degrees of evil are not considered by this account. If all evil characters are as evil as one another, we should intuitively reject the notion that a character like Satan is more evil than a character like Chad. Furthermore, our common intuitions also reject the notion that degrees of evil do not exist; even the terminology used by Haybron in his account supports the notion that degrees of evil exist and must be considered in any robust account.

### **5.5 — Evil people who are moved by the Good**

A robust account of evil character should not conflict with intuitive examples of evil people. Yet, Haybron claims that people who are moved by good desires to perform evil actions are merely misguided, and therefore not genuinely evil. However, in the second chapter of this project we argued that such a person in fact has an evil character, as seen in the case of the perverse benefactor. Haybron gives no supporting argument for thinking that the perverse benefactor is merely misguided, so let us reconsider the perverse benefactor and see if it is just merely misguided.<sup>25</sup>

As discussed in chapter two, the rationale behind a perverse benefactor’s actions can be explained as a self-caused delusion. David Jones, for instance, argues that Hitler, a paradigm of the perverse benefactor, was deluded by way of self-deception.<sup>26</sup> Jones also

argues that self-deceiving characters, like Hitler, are morally blameworthy for their self-deception and, therefore, are culpable for their evil actions as well.<sup>27</sup> If Hitler, and by extension the perverse benefactor, is morally blameworthy for his self-deception, then the claim that he was misguided or deluded does not morally excuse his reasoning. Since his reasoning is not morally excusable, and his reason for acting was altruistic, Hitler serves as an intuitively acceptable example of an evil person who is moved by the good.

Therefore Hitler—and by extension perverse benefactors generally—stands as a clear counter-example to Haybron's view of evil character and produces a criticism that cannot be easily dismissed. Such a clear counter-example to Haybron's account exists as further evidence that the consistency account is not a truly robust accounting of evil character.

## **5.6 — Summary**

Daniel Haybron's consistency account holds that a character is evil if she is wholly and consistently unmoved by the good. We argued in this chapter that Haybron's consistency account does not provide clear sufficient or necessary conditions for evil character to obtain. More specifically, the consistency account does not present a plausible distinction between the very bad and the truly evil. Haybron's view also implausibly implies that no evil characters actually exist. Haybron's view also denies the existence of degrees of evil character.<sup>28</sup> However, this denial is contradicted by the very way in which Haybron develops his theory. Haybron's unintentional use of degrees-of-evil terminology further demonstrates that such degrees are an intuitively acceptable and necessary component of theory of evil character. Finally, his view is unable to account for

evil people who are significantly moved by the good. This leads us to conclude that the consistency account is not an adequately robust accounting of evil character.

## Notes

---

- <sup>1</sup> Daniel Haybron, "Consistency of Character and the Character of Evil." in *Earth's Abominations: Philosophical Studies of Evil*, ed. Daniel Haybron. (Amsterdam: Rodopi, 2002), 63
- <sup>2</sup> Daniel Haybron, *Earth's Abominations*, 63 & 70
- <sup>3</sup> *Ibid.*, 63
- <sup>4</sup> *Ibid.*, 69
- <sup>5</sup> *Ibid.*, 69
- <sup>6</sup> *Ibid.*, 70
- <sup>7</sup> *Ibid.*, 70-71
- <sup>8</sup> *Ibid.*, 70 & 71
- <sup>9</sup> *Ibid.*, 70 & 71
- <sup>10</sup> *Ibid.*, 70
- <sup>11</sup> *Ibid.*, 70
- <sup>12</sup> *Ibid.*, 63
- <sup>13</sup> *Ibid.*, 70
- <sup>14</sup> Daniel Haybron, "Moral Monsters and Saints," *The Monist* 85, No. 2, 2002: p. 261
- <sup>15</sup> Luke Russell, "Dispositional accounts of evil personhood," *Philosophical Studies* 149, No 2, June 2010, p. 240-241
- <sup>16</sup> Barry, Peter Brian. "Moral Saints, Moral Monsters, and the Mirror Thesis." *American Philosophical Quarterly* 45 No.2, 2009, 169
- <sup>17</sup> Barry, *Moral Saints*, 169
- <sup>18</sup> Daniel Haybron, "Moral Monsters and Saints," *The Monist* 85, No. 2, 2002, 261
- <sup>19</sup> Daniel Haybron, "Consistency of Character and the Character of Evil." in *Earth's Abominations: Philosophical Studies of Evil*, ed. Daniel Haybron. (Amsterdam: Rodopi, 2002), 70
- <sup>20</sup> Daniel Haybron, *Earth's Abominations*, 70
- <sup>21</sup> *Ibid.*, 70-71
- <sup>22</sup> *Ibid.*, 70
- <sup>23</sup> *Ibid.*, 70
- <sup>24</sup> *Ibid.*, 69
- <sup>25</sup> *Ibid.*, 72
- <sup>26</sup> David H. Jones, *Moral Responsibility in the Holocaust: a study of in the ethics of character*, (Rowman & Littlefield. Oxford, Uk 1999), 137-142
- <sup>27</sup> David H. Jones, *Moral Responsibility*, 142
- <sup>28</sup> Daniel Haybron, "Moral Monsters and Saints," *The Monist* 85, No. 2, 2002, 261

## **Towards a Robust Theory of Evil Character: Chapter 6**

### **6.1 — Peter Brian Barry’s Extremity of Vice Account**

This chapter examines Peter Brian Barry’s extremity of vice account to determine whether or not it meets our criteria for a robust theory. When constructing his account, Barry was inspired by a metaphysical folk notion that his account attempts to explain.<sup>1</sup> He refers to this folk-notion as the mirror thesis, which he initially summarizes as follows:

“[A] person is evil just in case she is the perverse mirror image of the moral saint.”<sup>2</sup>

Barry begins development of his account by determining what constitutes a moral saint.<sup>3</sup> He adopts some suggestions from Susan Wolf’s work on moral sainthood to provide the basis for his account, although those portions of Wolf’s writings do not reflect her own final view.<sup>4</sup> At one point, Wolf suggests that moral saints are those persons who possess the most extreme virtues to the most extreme degree.<sup>5</sup> Consequently, Barry determines that evil persons must therefore be those who possess the most extreme vices to their most extreme degree.<sup>6</sup>

The definitions Barry uses for virtues and vices are integral to understanding this account of evil character. Barry notes that, according to one theory, virtues can be defined as “multitrack dispositional states,” which means that virtues, and correspondingly vices, “dispose their agents to act in certain ways in certain circumstances, to act for certain reasons with certain feelings.”<sup>7</sup> In layman’s terms, virtues and vices represent a person’s moral dispositions.

According to this view, if the moral saint is a person possessing the most extreme

virtues to their most extreme degree, then she is disposed to act extremely virtuously. Likewise, according to Barry, an evil character possesses the most extreme vices to their most extreme degree and is consequently disposed to act extremely viciously.<sup>8</sup> Thus, Barry formulates his account as follows: “an agent *a*,[sic] is an evil person just in case *a* possesses extremely vicious states of character.”<sup>9</sup>

Drawing on the same portions of Wolf’s analysis, Barry contends that extremity of vice exists in two different senses. Firstly, in order for a character to be evil, her vices must be the worst vices a person could have, e.g. cruelty and malice rather than laziness or cowardice, and secondly, those vices must be at their most extreme degree.<sup>10</sup>

The extremity of vice account intends to provide a firm distinction between merely very bad and truly evil characters.<sup>11</sup> His account also contends that there are no degrees of evil character.<sup>12</sup> Like Haybron and many others, Barry endorses the notion that the concept of evil ought to be applied only to the very worst possible sort of moral character.<sup>13</sup> There are very few people we could reasonably call moral saints, since those who possess the necessarily extreme virtues for sainthood are few and far between. Likewise, very few characters actually exhibit sufficiently extreme vices to be properly judged evil.

Barry explains that if a person exhibits extremely vicious character traits, her “moral well is poisoned.”<sup>14</sup> This poisoning renders any virtues one may possess morally irrelevant in comparison to the overwhelming extremity of one’s vices. This argument attempts to functionally explain how truly evil persons can occasionally behave morally well without diminishing their overall evil character.

From Wolf's initial writings on moral saints to his own account of evil character, Barry is able to formulate the final version of the mirror thesis, which states "a person is evil just in case he is a perverse reflection of the moral saint such that he suffers from extremely vicious character traits."<sup>15</sup>

## **6.2 — Lacking Degrees of Evil**

Barry's extremity of vice account is very nearly a robust account of evil character. Yet, some critical issues prevent it from being entirely satisfactory. The first criticism concerns a question of plausibility. Barry, like Haybron, endorses a degreeless view of evil character.<sup>16</sup> This creates a familiar problem for his account, since any theory that does not accommodate degrees of evil is intuitively implausible.

Despite Barry's endorsement of a degree-less view of evil character his account is surprisingly riddled with degrees-of-evil terminology. For instance, the idea that certain vices are more extreme than others is clearly dependent upon the existence of varying degrees of vicious extremity. As a corollary, in Barry's view, evil character is defined by the possession of extreme vices to extreme degrees. By extension, it should be fine to say that one is more evil than another if one is more vicious. Barry consistently makes statements that certain vices are greater or lesser than others, which suggests that his dismissal of degrees of evil is incoherent.

A proponent of Barry's view could respond that while Barry does use terminology that presupposes degrees, he does not leave any room for degrees of evil in his account. Barry defines an evil character as one having the most extreme vices to their most

extreme extent, and there is no place for any other degrees of extremity beyond that. In Barry's view, one cannot be more vicious than one who is most vicious. Thus, given Barry's definition of evil, it appears that one cannot be more evil than is required for evil.

Yet, this defense of Barry's view creates more problems than it solves as this means that only those who are the most vicious could ever be rightly considered evil. This is unintuitive, as not every evil character is as vicious as every other evil character—this is evinced by Haybron's intelligible comparison between Chad and Satan, as discussed in the previous chapter.<sup>17</sup>

A proponent of the extremity of vice account could attempt to insist that degrees of evil do not exist by turning to the reasoning presented by Daniel Haybron. As was made apparent earlier in the previous chapter, Haybron believes no real distinction is made when a character is described as evil, as opposed to very evil.<sup>18</sup> Similarly, a proponent of Barry's view could argue that once someone is sufficiently vicious to be evil, the extent to which her extremity of viciousness extends matters not at all.

However, envision being evil as being in a swimming pool, with the philosopher's task of defining evil character being like that of a lifeguard saving lives. It is possible for a person to drown in either the shallow or the deep end of the pool, and the outcome—death—is the same in either depth of water. Yet, for the lifeguard, the depth of the water they must dive into to save a drowning person matters significantly. Likewise, philosophers of evil ought to desire to be precise when defining degrees of evil.

### **6.3 – Too Narrow**



The next criticism is that the extremity of vice account is too narrow in scope. This point can be best illustrated by considering the perverse benefactor once more. As described in chapter two of this project, the perverse benefactor is an intuitively plausible evil character. Yet, according to Barry's account, a perverse benefactor is not an extremely vicious person and would not be considered evil. For Barry, any person who is primarily disposed to act by desires to do good is not motivated by the most extreme vices and is therefore not an evil character. Recall that, in Barry's view, virtues and vices are "multi-track dispositional states" that give people their moral dispositions to act in certain ways. The perverse benefactor acts from an altruistic disposition, and thus, is not a vicious person, in Barry's view. The example the perverse benefactor serves to illustrate is that Barry's failure to account for some types of evil character renders his account too narrow in scope. This criticism reminds us of Barry's criticism of Colin McGinn's account, that McGinn could not accommodate the concept of an evil person who did not have a reversed hedonic disposition.<sup>19</sup> Just as McGinn was not able to accommodate an evil person who did not possess a reversed hedonic disposition, so too is Barry unable to accommodate an evil person who is not extremely vicious.

Yet, with both criticisms considered, the extremity of vice account is very nearly a robust account. Barry's views could easily be altered to accommodate degrees of evil without sacrificing any of his account's explanatory power. Furthermore, he compellingly explains the dispositions and behaviors of moral monsters, who are actually extremely vicious. However, Barry's account precludes certain types of evil characters, like the perverse benefactor, from being evil. Given how well this account performs, its

suggestion to consider a person's virtues and vices— namely, moral dispositions—might be worth further exploration and will be discussed in the following chapter, which examines Luke Russell's dispositional account.

#### **6.4 — Summary**

Peter Brian Barry seeks to explain the metaphorical sense in which an evil character mirrors a moral saint. To do so, he developed a mirror thesis, which portrays the evil person as an opposite reflection of a moral saint. Barry believes that one analysis of moral sainthood that Susan Wolf constructed is appropriate, and so applied the mirror thesis to it, concluding that an evil person must be an agent who possesses extremely vicious character traits.<sup>20</sup> However, Barry fails to say why and in what way(s) an evil person ought to be the perverse reflection of a moral saint, which leads to criticisms of the extremity of vice account's narrow focus on extremely vicious people when it purports to encompass all types of evil character.<sup>21</sup> Additionally, Barry adheres to a popular, though unintuitive, notion that evil exists without degree or variance, despite the account's dependency upon degrees-of-evil terminology.<sup>22</sup>

A number of important lessons can be taken from our examination of this account. For one, we continue to see the importance of including degrees of evil in our account of evil character. We are also reminded that making our account too exclusive will render it intuitively implausible. Most importantly, however, are the positive elements present in Barry's view. Dispositions—in this case, virtues and vices—prove important for explanations of saintly and monstrous moral character. Furthermore, one lesson that can

be taken from Barry's account is that the manner in which a person is disposed to regard vicious actions is of greater importance than how a person is motivated to act.

## Notes

---

<sup>1</sup>Peter Brian Barry, "Moral Saints, Moral Monsters, and the Mirror Thesis." *American Philosophical Quarterly* 45 No.2, 2009, 163

<sup>2</sup> Barry, *Moral Saints*, 164

<sup>3</sup> *Ibid.*, 164-8

<sup>4</sup> *Ibid.*, 169-171

<sup>5</sup> *Ibid.*, 171

<sup>6</sup> *Ibid.*, 171

<sup>7</sup> *Ibid.*, 171

<sup>8</sup> *Ibid.*, 171

<sup>9</sup> *Ibid.*, 171

<sup>10</sup> *Ibid.*, 171-2

<sup>11</sup> *Ibid.*, 172

<sup>12</sup> *Ibid.*, 173

<sup>13</sup> *Ibid.*, 171-3

<sup>14</sup> *Ibid.*, 172

<sup>15</sup> *Ibid.*, 173

<sup>16</sup> *Ibid.*, 173

<sup>17</sup> See chapter five

<sup>18</sup> See chapter five

<sup>19</sup> *Ibid.*, 166

<sup>20</sup> *Ibid.*, 169-171

<sup>21</sup> *Ibid.*, 163

<sup>22</sup> *Ibid.*, 173

## Chapter 7: Luke Russell's Dispositional Account

### 7.1 — The Account

This chapter concerns Luke Russell's dispositional account, which is the last of the accounts of evil character that this project examines. Following the same format as the previous chapters, we will first present the dispositional account as Russell states it. Some criticisms will then be considered, which will demonstrate why this view is not an entirely robust accounting of evil character. After these criticisms, we will conclude with a summary of Russell's account and discuss the important lessons we have gleaned from its analysis, before we move on to the final chapter of this project.

Russell believes we have eight specific intuitions about evil persons or their character, and he seeks to create what he believes is the best account to explain these intuitions.<sup>1</sup> The eight intuitions Russell delineates are as follows: (1) Evil persons do actually exist. (2) Evil persons are rare. (3) Evil persons deserve our strongest moral condemnation. (4) The phrase "he did it because he was evil" is a meaningful statement. (5) Not every evil-doer is an evil person. (6) It is possible to become an evil person by performing evil actions. (7) It is possible for an evil person to become a good person. (8) Not every evil person performs, attempts, or intends evil actions.<sup>2</sup> We should consider these eight intuitions to be implicitly included in the criteria that must be met by any and every robust account of evil character.

Russell believes that a person is evil if and only if she is strongly and highly fixedly disposed to perform evil actions in conditions favorable to her autonomy.<sup>3</sup> Since

many conceptual elements comprise Russell's view, these must be clarified if one is to understand Russell's account correctly.

In Russell's vocabulary, fixity refers to the difficulty with which one's dispositions are able to change over time.<sup>4</sup> Remaining consistent with Russell's sixth and seventh intuitions, the dispositional account allows that some evil persons may become good persons, and vice versa. However, Russell believes that such transformations only rarely occur, and they are very difficult to achieve.<sup>5</sup> Therefore, to reflect the difficulty of changing one's entire disposition, Russell describes an evil person's character as highly fixed. The old adage "a leopard cannot change its spots" perhaps best reflects the inherent difficulty involved in overcoming and changing one's character.

When Russell describes evil character as strongly disposed to perform evil actions, he attempts to articulate that an evil person must be greatly inclined toward committing evil actions.<sup>6</sup> Russell acknowledges the fact that non-evil persons do occasionally cause great harm, but they do not have a great, or strong, propensity for such actions.<sup>7</sup> Evil people, according to Russell, are strongly inclined to commit evil actions to such an extent that it is contrary to their natures for them to do otherwise when in conditions that favor their autonomy. An evil person is strongly disposed to commit evil actions; to do otherwise is as contrary to her inclinations as it is for someone who strongly prefers chocolate ice cream to willingly choose vanilla instead.

Russell's view of dispositions is informed, in part, by the work of J.L. Mackie, who defines a disposition as a property possessed by things that causes them to X under certain conditions.<sup>8</sup> To demonstrate his claim Mackie provides an example, which states

that a piece of glass is fragile if it is disposed to break when struck. Similarly, for Russell, a person is evil if she is disposed to perform evil actions under certain conditions — more specifically, under conditions that favor one’s autonomy.<sup>9</sup> For Russell, conditions that favor autonomy include, but are not limited to, situations in which one is neither coerced, forced, tricked, nor presented with only one available course of action.<sup>10</sup> Russell notes that there have been cases in which people have “described themselves as being alienated from their [evil] actions in some way,” whether in coercive political environments or scenarios like the Milgram experiment, where subjects were directed to administer what they thought were increasingly painful shocks to people they believed were fellow participants in a study on learning.<sup>11</sup> In light of these cases, Russell excludes such people from his definition of evil character with the understanding that they would not perform evil actions if they were able to do “what they really wanted to do.”<sup>12</sup> To expand on this idea, he states that “[e]ven if it is true that a very large proportion of people are disposed to perform evil actions under some conditions, it is likely that only a small proportion of people are disposed to perform evil actions when they are able to do what they really want to do.”<sup>13</sup> Thus, Russell’s specification of the conditions under which evil actions are performed renders his account consistent with his second intuition, which suggests that evil persons are rare. To summarize this position, he states, “[One is evil] only if [one] is strongly [and highly fixedly] disposed to do evil even when [one] is not pushed in that direction by circumstance.”<sup>14</sup>

Russell asserts that evil actions can be understood according to one of two views: psychologically thick, or thin accounts of evil action.<sup>15</sup> According to the thin view, an evil

action is a culpable act that is “appropriately connected” to extreme undeserved harm.<sup>16</sup> The thick view uses a similar understanding, but additionally requires that the harmful act be performed by someone with a certain psychology.<sup>17</sup> Russell intends for his account to work with either conception, and he believes it functions equally well under either view.<sup>18</sup>

Finally, one should note Russell’s emphasis on the performance component of his analysis. According to Russell’s view, a person would not be evil if she were not disposed to actually perform evil actions. Essential to Russell’s account is the theory that evil characters are disposed towards the performance of real actions, suggesting that one’s mental state alone is not sufficient to define a character as evil in this view.

As Russell claims, his dispositional account conforms to each of the eight intuitions previously articulated, and it therefore appears to be a highly desirable account of evil character. Russell’s account is even accepting of the concept of degrees of evil, further increasing its desirability for the purposes of this project.<sup>19</sup> Yet, the dispositional account is not without some problems.

## **7.2— Unclear Elements**

As has been established in previous chapters, a robust account should be clear and precise in its definition of evil character.<sup>20</sup> However, Russell’s dispositional account tends to rely upon overly vague terms. For instance, even though readers should be able to understand easily what Russell means when he requires dispositions to be strong and fixed, these terms refer to qualities that cannot be plausibly justified. As mentioned previously, strength of disposition refers to how likely or inclined an individual is to

perform the action that she is disposed to do. The stronger one's disposition, or the greater one's inclination, the more likely an individual is to perform a certain act.

Similarly, Russell uses the term fixity to refer to the unlikelihood that one's dispositions will change over time. Whether the difficulties facing such change are meant to express the mental effort required to facilitate such a change, or merely the sheer improbability of a change occurring, is immaterial.

Although the components of evil character required by Russell are clearly articulated, the extent to which a disposition is sufficiently difficult, or unlikely, to change in order for a character to become evil remains problematically vague. The same can be said about exactly how strong a disposition must be. Like each of the previously examined authors, Russell needs to be more precise. For example, when referring to the strength of a disposition, Russell's account should discuss the extent to which a disposition is likely to become manifest in the conditions he specifies. Furthermore, he should also explain the reasoning that led him to specify some particular level of likelihood as sufficient. In addition, he should defend his position regarding how difficult it must be to change a disposition in order for that disposition to be a highly fixed one. Because these vague components of Russell's account require further explication, we have no certainty that they will not produce further problems that will need to be addressed. For example, Russell has stated that his view of dispositions allows them to be partial causes of their own manifestation; therefore, an overly strong disposition could be determined to become manifest, almost mechanistically, whenever appropriate conditions arise. In other words, just as a piece of glass is mechanistically determined to shatter



when struck with sufficient force, so too might a strongly disposed individual be mechanistically determined to perform an evil action when in autonomy-favoring conditions. If Russell believed that strength of disposition was necessary for his account, then it would seriously undermine our intuitions about free will and agency. This example of such an overly strong disposition is, of course, an exaggeration and is not meant to be a serious criticism of Russell's account. Rather, it is meant to demonstrate the manner in which various readers could mistakenly think that Russell holds any number of similarly problematic views, considering his vague and unqualified specifications regarding the elements of strength and fixity.

To strengthen his account, Russell should properly define, and justify, the elements he proposes are sufficient for evil character to obtain, as well as what would constitute their insufficient counterparts. For example, he should delineate the difference between a weak disposition and a regular, or strong, disposition in a character concretely. As stated in each previous chapter, a robust accounting of evil character should be precise and clear when establishing its conditions. Because Russell's account is vague in its explanations, we are unable to determine what specific values Russell has in mind, and therefore we are unable to stipulate whether or not the dispositional account provides even remotely adequate conditions for evil character to obtain.

### **7.3 — The Lazy**

Earlier in this project, we discussed one of Haybron's criticisms of Kekes's account. With the example of the malicious quadriplegic, Haybron demonstrated that

Kekes's account could not accommodate any evil person who was unable to perform evil actions.<sup>21</sup> A similar criticism can be made of Russell's theory by substituting laziness for quadriplegia. For example, a malicious, yet lazy, person would not be inclined to perform evil actions. Instead, such a character is more disposed to be sedentary. Following Haybron's example, this lazy character would perhaps enjoy hearing of, or seeing, the pain and death of her fellow humans, but it is likely that she would only do so by viewing the performance of evil actions on the internet from the comfort of her armchair. She would not be disposed toward doing anything, let alone taking up a hacksaw and going out to do evil herself. Suppose that this character's disposition towards sedentariness is very strong, highly fixed, and she retains such a disposition in conditions that favor her autonomy. Such a character would be an intuitively acceptable example of an evil character, yet, Russell's account would not consider her to be evil. Although this character is merely an example, people like our hypothetical character do exist, and they visit illegal websites to watch videos of torture, mutilation, and murder. These malicious lazy individuals present a direct counter-example to Russell's view, and since such a clear counter-example exists, Russell's account cannot be considered a completely robust account of evil character.

In reply to this counter-example, a proponent of Russell's view could argue that actively and willingly seeking out these illegal websites constitutes an evil action; if so, the character of the sort described above would in fact be disposed towards evil acts. While it is intuitively acceptable that actively seeking out such things could be considered an evil action—Todd Calder, for one, thinks so—this counter-example is not dependent

upon the proponent's definition.<sup>22</sup> For example, an exceptionally lazy malicious person could simply enjoy accidental observations of pain and death, such as those that might appear on the evening news broadcasts. In this case, the person is exposed to evil actions by only chance: she is disposed only to sit in her armchair and watch television, which is exactly what we would expect of an exceptionally lazy person. Such an extremely lazy character is perhaps an oddity, but certainly still plausible, and therefore stands as a counter-example to Russell's view. Therefore, we must look elsewhere for the locus of evil character, as the above counter-example shows that it is not found in one's disposition to perform evil actions.

#### **7.4 — Summary**

Russell's theory of evil character holds that an individual is evil if she is strongly and highly fixedly disposed, to perform evil actions when freely able to do so. Like the views discussed in previous chapters, Russell's theory contains some significant problems. Russell's definition of evil character is vague and requires a significant amount of explanation and justification before any judgments of its overall plausibility can be made. In addition, Russell's account faces a counter-example of an extremely malicious person who is not disposed towards actually performing evil actions. The existence of such a counter-example challenges the robustness of this account. Like the example of the malicious quadriplegic addressed in chapter three, the counter-example of the lazy person who enjoys nothing more than watching the suffering of others suggests that one is hardly made a better person simply because one is lazy and lacks a disposition to act.

Despite its faults, Russell's dispositional account is compelling. The idea that evil character is a combination of one's desires and one's propensity towards action is alluring to our intuitions.<sup>23</sup> However, Russell fails to recognize that it is not one's disposition to act that is paramount, but rather one's disposition towards the act. This will be discussed at length in the following, and final, chapter. We have now considered the manner in which action alone (Kekes), affect (McGinn & Haybron), and disposition (Barry & Russell) each missed the locus of evil character. However, none of the examined authors have presented a deliberative account, or, an account that considers the basis for evil character to be what a character *thinks* about evil actions, and what their rationale concerning such actions. The purpose of the following chapter is to conclude this project by presenting such an account.

## Notes

---

<sup>1</sup> Luke Russell, "Dispositional accounts of evil personhood," *Philosophical Studies* 149, No 2, June 2010, 232

<sup>2</sup> Luke Russell, *Philosophical Studies*, 232

<sup>3</sup> *Ibid.*, 247

<sup>4</sup> *Ibid.*, 244-5

<sup>5</sup> *Ibid.*, 244-5

<sup>6</sup> *Ibid.*, 234

<sup>7</sup> *Ibid.*, 234

<sup>8</sup> J.L. Mackie, "Dispositions, Grounds and Causes," *Synthese*, Vol. 34, No. 4, Dispositions (Apr., 1977), 361

<sup>9</sup> Luke Russell, "Dispositional accounts of evil personhood," *Philosophical Studies* 149, No 2, June 2010, 247

<sup>10</sup> Luke Russell, *Philosophical Studies*, 248

<sup>11</sup> *Ibid.*, 247 The Milgram experiment was conducted at Yale University in July 1961 by Stanley Milgram. Stanley Milgram, "Behavioral Study of Obedience," *Journal of Abnormal and Social Psychology* 67, No 4, 371-8

<sup>12</sup> *Ibid.*, 248

<sup>13</sup> *Ibid.*, 248

<sup>14</sup> *Ibid.*, 248

<sup>15</sup> *Ibid.*, 232

<sup>16</sup> *Ibid.*, 232

<sup>17</sup> *Ibid.*, 232

<sup>18</sup> *Ibid.*, 232

<sup>19</sup> *Ibid.*, 243

<sup>20</sup> See chapters three through seven, for the importance of clarity.

<sup>21</sup> See chapter three

---

<sup>22</sup> Todd Calder, "Evil and Its Opposite", *Journal of Value Inquiry* 49 (2015), 119

<sup>23</sup> Luke Russell, "Dispositional accounts of evil personhood," *Philosophical Studies* 149, No 2, June 2010, 248. I say that Russell's account includes an element of our desires as he uses our 'wants' to exemplify what constitutes an autonomy favoring condition.

## **Chapter 8: Towards a Robust Account**

### **8.1 — Chapter Structure**

This chapter forms the conclusion to our project. As such, it will summarize and explicate the various arguments that were presented in previous chapters. First, we will present a summary of the views held by the five authors this project examines, including the lessons that were gleaned from them. We will then conclude the project by sketching an account of evil character based upon those lessons. While this chapter will note the benefits and disadvantages of this sketch, further development will have to come in a future work. Once this chapter has completed these goals, the project will conclude with an exhortation that the philosophical community should implement the lessons developed here to inform future attempts at developing a definitive robust account of evil character.

### **8.2 — The Five Accounts**

We examined five different accounts of evil character, each of which was chosen to exemplify a specific type of account across a wide range of theories. Kekes's account is formulated entirely around a person's external actions, rendering psychological elements unimportant to his analysis.<sup>1</sup> Thus, his view is an excellent archetype of an external understanding of evil character. Kekes's view is also quite broad, as it does not consider how or why one frequently causes harm to be relevant to whether one's character is evil or not. This lack of internal exploration and its overly vague construction led to the formulation of some counter-examples, and Kekes's account was deemed

unrobust.<sup>2</sup>

McGinn's account is the opposite of Kekes's, for all intents and purposes. McGinn's model of evil character is entirely psychological and therefore provides an entirely internal understanding of evil character. This view disconnects itself from motivational components, making it an effective archetype of an entirely psychological account. McGinn's theory was also the first strict, or narrow, account we examined, since people require certain specific traits to be included in his view. However, this narrowness proved to be a problem, given the inevitable counter-examples. When combined with requirements that are difficult to discern, these counter-examples prove that this account is also not sufficiently robust.<sup>3</sup>

Haybron's theory is yet another example of an internal account, but its formulation is unique, since his broad view attempts to examine one's entire character in a dissimilar manner than the other four authors' approaches. However, his account still makes specific requirements for evil character to obtain. For instance, Haybron argues that evil characters must be unaffected positively by the good, a contention that ultimately limits his overall view. While Haybron's theory is able to offer an explanation of why evil characters perform evil actions, his account is extremely vague, making use of terms without providing adequate explanations of them. This shortage of clarity, along with a general lack of information concerning what evil characters are actually like, makes his account susceptible to some counter-examples. Overall, Haybron's account is unable to provide a robust theory due to its overly vague formulation.<sup>4</sup>

Barry's account serves as an archetype of the mirrored view of evil character. An

internal account that sets its conditions upon one's vices, Barry's theory easily explains the motivations of evil characters, yet he is able to do so only in a limited way. Although Barry's account functions remarkably well at explaining a narrow type of evil character—the moral monster—we are looking for an all-encompassing theory of evil character. Accordingly, there are counter-examples to Barry's views, as well as a problematic implementation of vague elements, such as his definition of extreme character traits. In short, Barry's account was quite close to being a robust account of evil character, but unfortunately it does not meet all the criteria.<sup>5</sup>

Finally, we examined Russell's account, yet another internal theory of evil character. However, Russell's view is unique when compared to the others because it, like Kekes's view, does not consider why an evil person has the disposition that they do to be relevant to the account. Russell's view states that an evil person is one who is strongly and highly fixedly disposed to perform evil actions when she is able to do as she wishes. This theory is appreciably sturdier than the previous four, as it is able to encompass examples that had posed problems for the previous views. Yet, it too is overly vague, and despite its success in generating a fairly intuitive view, it does still meet with a counter-example. Though not the definitive view, Russell's account stands out as the closest of the five accounts examined to meeting the criteria for a robust theory.<sup>6</sup>

Though we did not come across an entirely robust account of evil character, our examination of these five views allowed us to glean some valuable lessons. First, we learned that a robust account must have terms that are understandable and properly defined. Russell's and Barry's accounts are fairly successful in this regard.



Second, we learned that it is better to make the locus of evil character internal, rather than external. Kekes's account does not fare well because it puts the sole requirements for evil character on one's actions alone. Correspondingly, McGinn's account leans too far in the other direction and does not consider external behavior at all. Thus, it seems that the accounts that are most coherent are those that set the locus for evil internally while actively seeking to use the internal locus to explain the external behaviors of evil people.

Third, we learned that we must be careful not to make our account of evil character too narrow. Russell's theory fares the best overall, and it should not confront many counter-examples beyond the one portrayed in this project. Kekes's and Haybron's views are too narrow by ruling out evil characters who do not perform evil actions or who have a good side, respectively. Russell's and McGinn's views, and to a lesser extent Barry's, each are too narrow by hinging their accounts on single elements and asserting that individuals possessing these elements are evil. For Russell, that element is a person's dispositions, for McGinn that element is a person's hedonic orientation, while for Barry it is the possession of extreme vices. The problem with these accounts is that it is relatively easy to find counterexamples where an intuitively plausible evil character can be shown to lack the element touted as essential for evil character.

Fourthly, and perhaps most controversially, we learned that setting conditions—even hypothetical ones—that require characters to act can be detrimental to the development of a robust account. Russell's view, and to a lesser extent Kekes's as well, require evil characters to perform evil actions when they are freely able to do so. Yet,

both of these accounts face counter-examples in which a character intuitively appears to be evil but would not actually perform evil actions. While it certainly seems intuitive, and therefore correct, to assume that evil people are evildoers, it is clearly detrimental for an account to expect—even counter-factually—that an evil person must perform evil actions.

Lastly, and perhaps most unexpectedly, we learned that a robust account of evil character should not be based upon one's affects, or lack of affects.<sup>7</sup> It seems that since humans have such a wide and varying spectrum of feelings and affects, basing an account even partially upon one's affects could potentially impair that account's viability. Affects remain an important aspect of being human, but we should seriously consider leaving them out of definitions of evil.

### **8.3 — A Robust Sketch**

Having considered these five lessons gleaned from our analyses of accounts of evil, we are able to ask what would, in fact, constitute a robust account of evil character. For starters, a robust account should strictly adhere to providing an internally located definition of evil character, though such a definition would have to avoid including one's affects while still providing a functional avenue for those affects. In other words, a robust account should not include character affects, as far as is possible, but be formulated in such a way as to allow for a variety of affects. Furthermore, a robust account would have to use simple and intuitively understandable terms in its definition, while making sure the terms themselves are properly defined and the claims justified. In addition, a robust account must be intuitively plausible, make a clear distinction between bad and evil

characters, and accommodate the intuition that there are degrees of evil. Having established these bounds, we will sketch a potentially robust account of evil character.<sup>8</sup>

The Deliberative View: A person is evil if she would conclude that an evil action is morally acceptable when aware of all relevant facts relating to that action.<sup>9</sup>

This account attempts to exemplify all the lessons we have learned in this project while also adhering to the established criteria, though a deeper examination may later be required to demonstrate the account's advantages fully. Our examination of this account will follow the structures adhered to by previous chapters. First, we will dissect this account, which we will call the deliberative account, and analyze it. We will then compare it to the five authors' theories to show how it is an improvement over these accounts. We will also examine how it fares when confronted with the counter-examples that were used to critique those accounts. Lastly, we will discuss some of this account's disadvantages which, while not fatal to the robust nature of this view, could make it less desirable than an equally robust future alternative.

To conclude something means to determine by deliberation that X is the case and believe oneself justified in holding that conclusion. In other words, it means to feel epistemically justified in believing that X is the case—X, in this instance, being the belief that a particular action is morally acceptable when it is, in fact, evil. To think that something—such as X—is morally acceptable means to agree with X in a moral way,

which is to view X as morally neutral—neither good nor bad—or as morally good. In simpler terms, to find X morally acceptable means that X does not affect one's conscience in an intellectually repulsive manner—X, of course, still representing evil actions. Since this project is not concerned with what constitutes an evil action, the reader should feel free to substitute any reasonable definition of evil action for these purposes. For the sake of simplicity, the definition this project assumes, but does not endorse, is similar to the one put forward by Kekes, which is: any action that causes significant undeserved physical or mental harm.

When the deliberative account refers to an evil character as one that concludes evil actions to be morally acceptable when aware of all relevant facts, it emphasizes the notion that some people tend to delude themselves about some facts, as was argued by Mike Martin and David Jones in their discussion of self-deception.<sup>10</sup> For example, Hitler must have been aware on some level of the fact that European Jews were not to blame for Germany's woes, but he deluded himself about that fact in order to permit his preferred belief that Germany was blameless. To explicate such an example, a character must conclude X to be morally acceptable after being aware of all the relevant facts pertaining to X, whether or not that character dismisses any facts through self-delusion, or otherwise. Any dismissal of the relevant facts indicates that one is not considering—or deliberating with—all the facts in a straightforward manner, it does not indicate that one is no longer considering those facts at all, or that one does not possess them. So, such a person is still culpably concluding that X is morally acceptable, even if one is not considering all relevant facts straightforwardly. Alternately, individuals who are not

culpable for their lack of awareness of some relevant facts regarding X are exculpated from being considered evil, since they were not aware of some of the facts. Therefore, a person is evil if she thinks herself epistemically justified in believing that an action that causes significant undeserved harm is morally good—or is neither good nor bad—when she has all the relevant facts.

This account does not include as possible moral agents very young children, plants, animals, or the mentally ill, insofar as they are unable to deliberate—i.e., think about—the facts appropriately. On the other hand, those children or mentally ill people who are capable of deliberating the facts appropriately are correspondingly capable of being considered evil by this theory. In addition, the deliberative view easily accounts for self-deceiving characters like the perverse benefactor, as the example of Hitler discussed above shows. Furthermore, the account also accommodates Haybron's malicious quadriplegic, as well as the malicious-lazy-person variant, as they are not morally repulsed by the pain and suffering they both relish. On the deliberative account, their lack of conscientious repulsion implies that they do not find anything morally objectionable about what they relish.

The deliberative view also takes into consideration each of the problems facing the five author's accounts. In response to such criticisms, this account is clear in its definition, simply declaring that concluding any evil actions morally acceptable is sufficient for evil character. This includes instances in which evil characters have compartmentalized their lives, as, for example, some Nazis did: they found nothing wrong with executing undesirables all day, and yet appeared to be loving, moral people in

their personal lives.<sup>11</sup> Like Russell's account, the deliberative view does not include those who are forced to do evil in situations like the Milgram scenario, since these characters do not find the evil actions they are performing morally acceptable. This view also maintains a distinction between the very bad and the truly evil by stating that one must find evil actions morally acceptable to be evil. According to this understanding, there can be many evildoers, but it is unlikely that many evil people exist, since most who commit evil actions do not believe them to be morally acceptable. This component also makes sense of the phenomenon Kekes noticed—that the amount of evil present in the world is much greater than the number of evil people—which led him to mistakenly adopt a hard reaction to evil character.

Since the deliberative account does not attempt to stipulate how, or why, evil people arrive at the conclusions that they do, it remains able to encompass all possible evil characters, which is something all of the previously considered accounts hoped, but failed, to achieve. The deliberative view is somewhat similar to the dispositional view as it concerns itself with how one is inclined to think, rather than how one is inclined to act. Finally, the deliberative view also explains why characters are motivated to perform evil actions without restricting evil actions to a narrow range of motivations. According to this view, an evil character must find evil actions morally acceptable, and if one is evil, one has no internal (psychological) moral reason not to perform evil actions.<sup>12</sup> Therefore, if one is not cowardly, lazy, physically impaired, or otherwise cognitively inhibited, one has no effective conscience repulsing oneself—in an intellectual sense—from committing evil acts. Ergo, the deliberative view allows evil to be performed for many different

reasons.

Most importantly, this account adheres to the established criteria for a robust theory. It provides a sufficient condition that appears to be compatible with all of our intuitions, as evinced by comparisons with the previous authors' accounts. To put it briefly, the sufficient condition in this account is that an evil person concludes that evil actions are morally acceptable. Furthermore, this also implies a necessary condition for evil character—namely, that one lacks an effective conscience. The deliberative view also permits degrees of evil in several ways. One's status as more or less evil can be determined by whether one finds it easy or difficult to consider evil actions morally acceptable. On the other hand, degrees of evil could depend on the severity of evil actions, since certain lesser evil characters might find the least severe actions morally acceptable, but may find more severe ones less palatable. This last point could explain, for example, why some evil people only steal but do not kill or torture, whereas other people are more liable to kill or torture than they are to steal. Thus, even though this account is only a brief sketch, we are able to see how it is an intuitively plausible and robust theory of evil character.

This sketch is not meant to be a definitive account of evil character. It is my hope that, in the future, I will see a truly definitive robust account of evil character developed that is based upon the criteria established in this project.

## Notes

---

<sup>1</sup> John Kekes, *The Roots of Evil*, (Ithaca NY: Cornell University Press, 2005), 2

<sup>2</sup> See chapter three

---

<sup>3</sup> See chapter four

<sup>4</sup> See chapter five

<sup>5</sup> See Chapter six

<sup>6</sup> See Chapter seven

<sup>7</sup> See Chapter's 4 and 5 for how McGinn's and Haybron's accounts contain affective elements to their own detriment.

<sup>8</sup> This account is largely inspired by the view put forward by Laurence Thomas in his book *Vessels of Evil: American Slavery and the Holocaust* (1993). Though this account has few similarities to Thomas's view now, it is important to credit him for the original structure that this account derives from.

<sup>9</sup> I am not committed to using the word 'conclude' to represent what I mean; which is that a person must come to believe that evil actions are morally acceptable from reasons. Neither am I committed to calling this type of view the 'deliberative view' as use of the word 'deliberative' could be confusing to some.

<sup>10</sup> David H. Jones, *Moral Responsibility in the Holocaust: a study of in the ethics of character*, (Rowman & Littlefield. Oxford, Uk 1999), 79-81

<sup>11</sup> Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil*, (New York: Penguin Books, 1964) As argued by Arendt

<sup>12</sup> A person has no moral reason not to perform evil actions in a similar sense to Laurence Thomas's more affect-based approach. Laurence states that evil characters lack inhibitions against performing evil actions because their consciences are not affected by the evil of their intended actions. I am contending that there is a more cognitive or psychological sense in which we can conceive their lack of affect than a purely affective one.



## Bibliography

- Ammer, Christine. *The American Heritage Dictionary of Idioms*. Houghton Mifflin, 1997
- Arendt, Hannah. *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: Penguin Books, 1964
- Barry, Peter Brian. "Moral Saints, Moral Monsters, and the Mirror Thesis." *American Philosophical Quarterly* 45 No.2, (2009): 163-176
- Calder, Todd. "Evil and Its Opposite." *Journal of Value Inquiry* 49 (2015): 113-130
- . "The Apparent Banality of Evil." *Journal of Social Philosophy* 34 No.3, (2003): 364-376
- Frankfurt, Harry G. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy* Vol 68 No 1 (1971): 5-21
- Haybron, Daniel. "Moral Monsters and Saints." *The Monist* 85, No. 2, (2002): 260-284
- . "Consistency of Character and the Character of Evil." in *Earth's Abominations: Philosophical Studies of Evil*, ed. Daniel Haybron, 63-78. Amsterdam: Rodopi, 2002
- . "Evil Characters." *American Philosophical Quarterly* 36 No. 2, (1999): 131-147
- Jones, David H. *Moral Responsibility in the Holocaust: a study of in the ethics of character*. Rowman & Littlefield, Oxford, Uk 1999
- Kekes, John. *Facing Evil*. West Sussex, Chichester: Princeton University Press, 1990
- . "The Reflexivity of Evil." *Social Philosophy and Policy* Vol 15 No 1, (1998): 216-232
- . *The Roots of Evil*. Ithaca NY: Cornell University Press, 2005.
- Mackie, J.L. "Dispositions, Grounds and Causes." *Synthese*, Vol. 34, No. 4, (1977): 361-370
- Martin, Mike W. *Self-Deception and Morality*. Lawrence: University Press of Kansas, 1986
- McGinn, Colin. *Ethics, Evil and Fiction*. 61-91. Oxford: Clarendon Press, 1997
- Russell, Luke. "Dispositional accounts of evil personhood." *Philosophical Studies* 149, No 2, (2010): 231-250
- Thomas, Laurence. *Vessels of Evil: American Slavery and the Holocaust*. Temple University Press. Philadelphia PA, 1993
- Wolf, Susan. "Moral Saints." *The Journal of Philosophy* Vol 79 No 8, (1982): 419-439