

ADDRESSING CRIMES OF PASSION WITH THE DEEP-SELF VIEW OF MORAL RESPONSIBILITY



ELIJAH PARISH

ABSTRACT

In this paper, I summarize and object to the “deep-self” view of moral responsibility as laid out by Susan Wolf in “Sanity and the Metaphysics of Responsibility.” My objection centers on how our intuitions regarding crimes of passion conflict with the conclusions drawn by the deep-self view. I then proceed to sketch out three possible responses which can be made by an adherent to the deep-self view and make my recommendations on how such adherents should proceed in further understanding moral responsibility.



I. INTRODUCTION

In her paper, “Sanity and the Metaphysics of Responsibility,” Susan Wolf draws similarities between multiple prominent theories of moral responsibility in the twentieth century. From these connections, she devises what she calls the “deep-self” view of moral responsibility. What unites these views is the intuition that we are morally responsible for our actions, both that are willfully caused by us, and that intentionally flow from some more fundamental part of ourselves.¹ In this paper, I sketch out some of the historical manifestations of the deep-self view from which Wolf draws, present a new objection regarding crimes of passion that is not discussed in Wolf’s paper, discuss three possible responses to this objection, and outline future work to be done on clarifying and formulating the deep-self view of moral responsibility.

II. CRIMES OF PASSION AND THEIR NATURE

Before continuing in the discussion of crimes of passion, it is important to make some clarifications about their nature, as well as the nature of responsibility. By crime of passion, I am referring to any socially undesirable behavior done without premeditation and motivated by intense emotion, particularly one that may be described as uncharacteristic, regardless of scale; this can range from something as extreme as murder or assault to something far more mundane, such as cheating on a partner. While often used in a legal context, I will be discussing crimes of passion in terms of a moral responsibility rather than a legal one. Robin Zheng, in her discussion of implicit bias, describes this responsibility as the distinction between attributability (being an expression of agency and inviting praise or blame) and accountability (being responsible for the social ramifications).² I refer to the former when I invoke moral responsibility. Zheng’s discussion of attributability is similar to Wolf’s discussion of the deep-self, insofar as she discusses an action where one can be held morally (rather than merely legally or socially) responsible as they are “distinctively subject to self-reflective awareness.”³ An example Zheng gives of this distinction is of a car crash: someone can accidentally cause damage to another person’s car, and

while they would not be blamed for this accident, they would still be financially responsible for the damages.⁴

III. VERSIONS OF THE DEEP-SELF VIEW

Returning to the deep-self views of moral responsibility, under Harry Frankfurt’s view, moral responsibility requires more than freedom from external coercion.⁵ What is needed for moral responsibility is for one’s actions to align with their second-order desires—often phrased as what one “wants to want” and what Frankfurt freely calls “willing.”⁶ These desires of the second order are in contrast to the first-order desires of merely wanting something.⁷ Moral responsibility, according to Frankfurt, derives from the fact that humans care about their desires. If someone both wants to want X and wants X, then they are responsible for X; but if someone wants X but wishes they do not want X, then they are not responsible for X. Frankfurt uses an “unwilling addict” as an example to elucidate this principle. The unwilling addict simultaneously wants and does not want to take the drug. What is critical is that the unwilling addict is not neutral between these two desires—they want to not want to take the drug.⁸ Because of this second-order desire to not take the drug, the unwilling addict cannot be considered morally culpable when they are physiologically compelled to indulge their addiction, as one can only be held responsible for acting in accordance with one’s second-order desires.⁹

Gary Watson’s view is similar. Rather than discussing first and second-order desires, he categorizes our desires into “mere” desires, ones we are stuck with as a consequence of being an organic being thrown into the world, and values, which express some deliberative judgment.¹⁰ For example, while someone may have an inborn desire for hedonistic pleasures, they might also have a value for managing and limiting their indulgent behaviors. Subsequently, according to Watson, this person is morally culpable for the actions that proceed from evaluations.

Charles Taylor, quite similarly, argues that what makes humans free agents is the ability to reflect upon themselves.¹¹ He suggests that if one’s character were beyond their control (determined by an outside force

1 Susan Wolf, “Sanity and the Metaphysics of Responsibility,” in *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, ed. Ferdinand Schoeman (Cambridge: Cambridge University Press, 1988), 49.

2 Robin Zheng, “Attributability, Accountability, and Implicit Bias,” in *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, ed. Michael Brownstein and Jennifer Saul (Cambridge: Cambridge University Press, 2016), 62–3.

3 Zheng, “Implicit Bias,” 64.

4 Zheng, “Implicit Bias,” 66.

5 Harry Frankfurt, “Freedom of the Will and the Concept of a Person,” *Journal of Philosophy* 68, no. 1 (1971): 14.

6 Frankfurt, “Freedom of the Will,” 6.

7 Frankfurt, “Freedom of the Will,” 6.

8 Frankfurt, “Freedom of the Will,” 12.

9 Frankfurt, “Freedom of the Will,” 12.

10 Gary Watson, “Free Agency,” *The Journal of Philosophy* 72, no. 8 (1975): 208–9.

11 Wolf, “Metaphysics of Responsibility,” 49.



such as a deity and thus wholly untouchable), then one would be a mere vehicle for causal forces.¹² According to Taylor, one can recognize their current character and ideals, thereby making their own flaws apparent. Subsequently, one can resist their natural inclinations and cultivate new habits, such as changing their environment. Due to these capacities, Taylor suggests that humans are free moral agents.¹³

All three of these deep-self views hinge on an individual's capacity to reflect on their own character.¹⁴ Each is guided by the insight that, in terms of moral responsibility, it is insufficient that one is merely the cause of their actions, but that one's actions are, in some sense, an expression of their character. These deep-self views explain why kleptomaniacs and victims of brainwashing are not morally responsible for their actions, as well as why adult humans are morally responsible, but animals and infants are not. All of these actors—even if they can act freely, as Frankfurt would put it—lack the requisite faculties to be held morally responsible: the second-order desires, the ability to follow their desires, the ability to deliberate and generate values, or the capacity for self-revision and self-reflection. While these views do not assuage all of the deterministic fears of metaphysical responsibility, they do establish “all the freedom it is possible to desire or to conceive.”¹⁵ Wolf uses her paper to address the issue that all of these theories assume that metacognition is not a value-neutral endeavor. However, there is another flaw that even her sane deep-self view shares with its predecessors.

IV. TENSION BETWEEN THE DEEP-SELF VIEW AND CRIMES OF PASSION

The deep-self views struggle to account for crimes of passion. Under any conception of the deep-self view, it does not seem reasonable to hold someone responsible for a crime of passion. If we were to use Watson's version, a crime of passion would be considered a mere desire that one is stuck with, rather than a desire that flows from one's values. Similarly, under Frankfurt's framework, it is reasonable to think of a crime of passion as an action performed by an individual lacking second-order desires, or the capacity to obey such desires.¹⁶ In either case, they are a being that is not morally responsible. Following Taylor's theory that someone who acts spontaneously does not have the opportunity to reflect and thus cannot be held responsible, a crime of passion is not

an action for which one can be held accountable. This may seem like an unacceptable conclusion. It does not seem reasonable to say that someone who loses their temper and hurts someone else is not morally responsible simply because “they were not thinking straight” or that “it was not really them.”

I see three possible responses to this objection: we may throw out the deep-self view entirely, we may accept that one cannot be responsible for crimes of passion, or we must find a way to reconcile crimes of passion with the deep-self view. It is the last of these that I want to focus on, as I do not wish to abandon the project of moral responsibility, yet I believe that concluding that one cannot be responsible for crimes of passion is an unacceptable conclusion about the narrowness of moral responsibility. First, I would like to share a few thoughts on the first two options. In favor of rejecting the deep-self view, we may question the grounds on which we ontologically privilege “deep” desires over “mere” desires. We may feel compelled to reject the deep-self view as a false anthropology—perhaps a vestige of soul theory—in favor of a more Nietzschean view: that man is an assemblage of competing drives without a privileged ego-consciousness, where no part of the self is more core than any other.¹⁷ On this view, we could say we are responsible for all of our drives, or we could say that we are responsible for none of our drives. The key is that we abolish the hierarchy according to which some drives are more core to our identities than others. A proponent of this view might say that it is conceited of us to think that deliberative actions are any more an expression of ourselves than impulsive ones and might be disinterested in any post-hoc rationalizations that could be offered to explain why one is not actually responsible for the drives of which they disapprove.

Coming to an agreement that one cannot be held morally responsible for crimes of passion is another possible response, where one cannot be held morally responsible, only legally responsible.¹⁸ One could consider a crime of passion as analogous to temporary insanity, putting it among the ranks of other “excusing conditions,” such as acting unintentionally, under coercion or with an altered state of mind.¹⁹ While plausible, I believe that this position renders the scope of moral responsibility too narrow and that a rush of emotion should not be considered an “altered state of mind” comparable to the effects of drugs.

Our final option is then to attempt to reconcile these conflicting intuitions. On the one hand, we are only responsible for things that flow

12 Wolf, “Metaphysics of Responsibility,” 49.

13 Wolf, “Metaphysics of Responsibility,” 49.

14 Wolf, “Metaphysics of Responsibility,” 49.

15 Frankfurt, “Freedom of the Will,” 17.

16 Frankfurt, “Freedom of the Will,” 11.

17 Friedrich Nietzsche. *Human, All Too Human*, trans. Helen Zimmern (Edinburgh: T. N. Foulis, 1910), 107.

18 Zheng, “Implicit Bias,” 64.

19 Zheng, “Implicit Bias,” 65.



from our deep-self, and on the other, we are responsible for actions we perform on a whim. I will attempt to reestablish moral responsibility for momentary outbursts by reflecting on the nature of metacognition and introducing some cognitive science. According to the psychological model of the dual-processing theory of cognition, we have two manners in which we make decisions. System 1 thinking is intuitive.²⁰ It is fast, automatic, emotional, and subconscious. This is also the system employed during crimes of passion. System 2 thinking, by contrast, is deliberative.²¹ It is slow, logical, and methodical. There are clear parallels here to Watson's position that one is responsible for deliberations, but not mere desires. When advocates of various deep-self views conceptualized moral responsibility, this is largely the thought process they had in mind. It is easy to see how one could be responsible for something they spend significant amounts of time reflecting on and deliberating, but it is less clear where the responsibility lies in System 1.

V. RECONCILING RESPONSIBILITY WITH INTUITION

One way we may go about reconciling the deep-self view with crimes of passion is by establishing a duty to "think straight." We may imagine someone who was caught cheating telling their partner that they were not "thinking straight." To this, the partner could respond, "Well, you should have been thinking straight." If we could establish a duty to "have been thinking straight," we may be able to resolve this issue of moral responsibility. But does such a duty exist? Is one morally required to engage their System 2 processing at specific times? We may agree with the cheated-on partner that there is some duty to think straight, but there are some problems with trying to establish such a duty from the original position that one is responsible only for their deliberations. A duty to think straight would entail an imperative way to know when to use System 2 thinking. This way, we could maintain that someone is not responsible for their immediate intuitions, but instead argue that they are only responsible for knowing when to override their intuitions and think deliberately.

The primary concern with such a duty is that the overwhelming majority of one's thinking is preconscious, or System 1 thinking. Object perception, immediate effect, and language generation, just to name a few, are all examples of your brain on autopilot. More importantly, knowing when and where not to engage System 2 thinking is a System 1 faculty. That is to say, knowing when to deliberate is an intuitive decision.

20 Jonathan St B.T. Evans, "In Two Minds: Dual-Process Accounts of Reasoning," *Trends in Cognitive Sciences* 7, no. 10 (2003): 454.

21 Evans, "In Two Minds," 454.

To say that we are responsible for knowing when to deliberate violates our original premise that we are only responsible for our deliberative thought processes. This is not to say that there is not a duty to think straight, it is merely to say that such a duty cannot be derived from our original deep-self view.

This means that if one is responsible for crimes of passion, there must be times (though preferably not all times, lest we lose the explanatory power of the deep-self view entirely) when they can be held responsible for their intuitive thinking. I previously mentioned that, according to Wolf, the guiding intuition of the deep-self view is that in order to be morally responsible for an action, one is not merely the cause of such an action, but such an action is an expression of one's deep-self. For one to be responsible for crimes of passion, there must be times when intuitive thought processes are expressions of character, and thus something for which one can be held responsible. I believe this to be the case. As Taylor points out, we are capable of reflecting upon and revising our character.²² Metacognition is a necessary factor in moral responsibility. While it is true that System 1 thinking is intuitive and subconscious, that does not mean it exists separately from our characters. A good analogy would be to compare System 1 to wearing glasses and System 2 to taking off and inspecting those glasses. Through inspecting your glasses, you can clean them, change the tint, replace the lenses, etc. Analogously, through metacognition, one internalizes one's beliefs and modifies one's own cognitive machinery. Because of this, deep-self is integrated into subconscious actions. Therefore, automated actions are a (partial) reflection of one's deep-self. An example to elucidate this is to consider psychoanalysis (which sought to understand the unconscious through free association), Freudian slips, dream analysis, etc. If we accept that the unconscious or intuitive mind is not an expression of character, then we would also be forced to accept that the entire psychoanalytic project is somewhere between misguided and absurd for looking at something completely distinct from one's character.

VI. THE COST OF RECONCILIATION

There is one immediate objection I would like to address, and through doing so I would like to outline possible future work on the deep-self view. The objection is that the reconciliation of the deep-self view with crimes of passion proves too much. A significant part of the elegance of the deep-self view is that it makes a distinction, according to which we are not responsible for transient or peripheral actions but

22 Wolf, "Metaphysics of Responsibility," 49.



only ones which, in a more fundamental sense, express our character or deep-self. However, by arguing that we are sometimes responsible for our intuitive processes, I have flattened the distinction between the deep and shallow-self, and therefore I have eliminated the deep-self view.

There are a few possible ways of responding to this concern, and I will touch on each of them. First, a more moderate version of Watson's original claim (that deliberation, not intuition, expresses character) could be made. Instead, we could say that deliberation provides a more direct access to character whereas intuitions provide indirect access, and thus the responsibility is tempered, maintaining the deep/shallow-self hierarchy. Alternatively, we could take refuge in the fact that we maintain one of the original deep-self intuitions that children, animals, kleptomaniacs, people subject to hypnosis or brainwashing, etc., are not morally responsible. This is because their cognitive machinery has been infected or co-opted by some outside force, or is merely underdeveloped, resulting in their actions not being an expression of character.

To demonstrate this distinction, I will return to Frankfurt's earlier example of the unwilling addict. Suppose once more that there is an unwilling addict with a physiological compulsion to a particular substance. When presented with such a drug, they only perceive one option: consumption. As outside agents, we know there are plenty of other options, but as their world is perceived, the addict is left with only one option. Subsequently, they cannot be faulted for not taking some other course of action, as their mind has inhibited them from seeing such alternatives. They cannot be held responsible for something beyond their control. However, if they were to be rid of their compulsion, they would be able to recognize a whole array of possible courses of action. In the first case of the compulsion, their lack of agency prevents their actions from being an expression of character by way of only one choice being present. Whereas in the latter case, due to the abundance of possible decisions, the addict can be held morally responsible.

One could object to this picture, suggesting that being overtaken with strong emotion is tantamount to a compulsion. Such a position would be in line with the previous discussion of saying that one is not responsible for crimes of passion. However, we may have a reason for wanting to say that there is a significant difference between actions arising from strong emotions and compulsions. It seems that when I have a sudden upswell in emotion, I can identify it, label it, and consider its causes; though I may still feel the emotion, I can moderate its effects significantly through this analytic process in a way I may not be able to do for an addiction, which I could be able to recognize and consider, but be less capable of mitigating.

It seems that one of the next steps for proponents of the deep-self view is further clarifying the distinction between emotion, compulsion and addiction, and character. For example, what makes a repeated action a compulsion and the other a feature of one's character? While we may be satisfied to say someone is not responsible for their addiction, to say that someone is not responsible for having anger issues because they wish they were not so angry could be dissatisfying. Moreover, proponents should continue to flesh out this distinction between emotion-driven actions and addiction-driven actions. I have begun to sketch out some thoughts, but more needs to be done to unite philosophical and psychological literature in order to further validate or rebut the deep-self view of moral responsibility.

While I appreciate that both rejecting the deep-self view and biting the bullet on crimes of passion are plausible views, I believe that a deep-self view that can accommodate crimes of passion is the best way forward. That said, I also recognize that this view is not without flaws. As I mentioned, the view requires a further understanding of emotion, compulsion and addiction, and character. Perhaps even more challenging, the view asks us to question many long-held philosophical prejudices about the scope of rationality. Rather than maintaining a strict dichotomy between rationality and emotion, or body and spirit, this position asks us to view the human being as having beliefs, intuitions, and instincts all integrated within oneself. I hope that recognizing this broader and more integrated picture of the human being can also help paint a clearer picture of the issue of moral responsibility.





Elijah Parish is a sophomore studying philosophy, psychology, politics, and economics at the University of North Carolina. He is particularly interested in ethics, political philosophy, and the problem of Nihilism. In his spare time, he enjoys reading, writing, and playing Dungeons & Dragons. His preferred means of procrastination is listening to podcasts at two-times speed.

DOI: 10.33043/S.16.1.26-37

