

Melissa Kwan

Prof. Bernhard Nickel

Philosophy 3

13 December 2018

*Determinism* is the theory that the states of the universe inevitably proceed in one way, bestowing no control on any agent. The philosophical argument known as *The Clash* deduces the troubling implications of such a worldview. The argument goes as follows: If determinism is true, all events and outcomes are fixed, destined to occur based on the previous state of events. It follows that no being, rational or otherwise, can act in a way that leads to a state of affairs other than the one that actually occurs. The second part of the argument relies on the “could have done otherwise” premise, which states that agents are not free unless they could have acted differently than the way they did. Combining those two premises, The Clash argument holds that no actions are free. But if no action is free and freedom is a requirement for moral responsibility, The Clash concludes, no agent can be morally responsible for anything — we cannot reconcile determinism with moral responsibility.

Dissatisfied with this conclusion, compatibilists such as Susan Wolf seek to refute The Clash by undermining its premises. Wolf specifically targets the “could have done otherwise” premise, which she counters with an alternative conception of freedom: the Sane Deep Self.<sup>1</sup> The deep self condition stipulates that an agent is free if she performs the action in accordance with her superficial desires, which in turn must be under the control of her Deep Self. Thus, it is not enough for an agent to simply *desire* to do something; she must have some capacity to change

---

<sup>1</sup> Wolf 153.

and control those superficial desires. While Wolf does not explicitly characterize the Deep Self, she defines it loosely as a synthesis of Frankfurt, Watson, and Taylor's interpretations of fundamental selves. However, as Wolf points out, one's underlying values are ultimately still not under our control, and it is possible for one to be absolved of moral responsibility if one's moral beliefs are irrevocably wrong.<sup>2</sup> Thus, she argues that the deep self condition alone is necessary but not sufficient grounds for freedom.<sup>3</sup> Here is where Wolf differs from her contemporaries: she introduces the condition of sanity, which stipulates that the agent who performs the action must have moral knowledge grounded in reality. To be sane, an agent must have the ability to both "cognitively" and "normatively" adjust her behavior based on correct conceptions of right and wrong.<sup>4</sup> The sanity condition is Wolf's way of closing the sufficiency gap that the Deep Self condition leaves, which completes her argument that moral responsibility is compatible with determinism.

**In this paper, I will defend Wolf's refutation of the conclusions of The Clash and revise her conditions of freedom: to be free, agent must be fundamentally sane and able to modify both her desires and actions based on the different incentives.** I will begin by verifying that Wolf's alternative conception of freedom successfully invalidates the conclusions of The Clash. I will then present a revised necessary condition — the ability to adapt one's *desires* based on incentives — that constitutes a more nuanced conception of free agency. By challenging Wolf's necessary conditions, I will present a broader interpretation of freedom that allows for a wider scope of moral responsibility.

### Section 1: Wolf's Argument Against The Clash

---

<sup>2</sup> Ibid., 152.

<sup>3</sup> Ibid., 155.

<sup>4</sup> Ibid., 158.

Because the conclusion of The Clash states that determinism is altogether incompatible with moral responsibility, Wolf can refute its conclusions with a counterexample. Given the Clash's logic that agents are morally responsible for their actions if and only if they are free, we can deduce that an agent who is morally responsible is free. Consider a person who bakes a cake for a friend. She wants to take that action, and her deep self has determined that it supports her superficial desire to take that action — perhaps she wants to be the type of person who does kind things for others. Furthermore, her deep self is both sentient and sapient: her conception of the world, both in the physical sense and the moral sense, is tethered to reality. Because the action of baking a cake is morally consistent with her character and she has an accurate knowledge of right and wrong, her character can be morally evaluated based on that action. Given that morally responsible agents are necessarily free, we can conclude that our agent who meets Wolf's Sane Deep Self condition is in fact free.

## Section 2: Modifying the Conditions for Freedom

I will now add a stipulation to Wolf's definition of freedom to account for cases of willpower and self control. To do so, I will examine both an extreme alcoholic and a person who is simply prone to procrastination. The alcoholic feels compelled to drink despite not wanting to be the type of person who drinks; the procrastinator procrastinates despite wanting to finish her projects early. To simplify the case, assume that the alcoholic was born with her addiction and therefore cannot be held morally responsible for allowing her condition to arise. In both cases, the agents' actions match their superficial desires but clash with their deeper desires. However, they differ in their moral responsibility: the extreme alcoholic is suffering from addiction and therefore not responsible for her actions, yet the procrastinator is responsible for the

consequences of poor time management. To justify why these two agents deserve differing moral treatment, we must delve into differences in our reactive attitudes. In the alcoholism case, we instinctively absolve the agent of moral responsibility due to her inability to control her first-order desires. Perhaps one could also say that the procrastinator is unable to control her first-order desires, but we instinctively view the second agent as responsible for her “lack of self-control.” While it is true that in cases of addictions or other extreme compulsions, agents may truly lack the deep-self governance to be considered free, there are a multitude of more moderate “willpower” cases that the Deep Self condition fails to address.

To address this distinction, we must examine the differences in how the alcoholic and the procrastinator are free. It is difficult to characterize the difference between total absolution and moral responsibility, between being truly unable to control one’s desires and being simply melodramatic. To do so we might consider Van-Inwagen’s depiction of compatibilism, which loosely states that an action is free if given the same state of the world and different incentives, the agent would have behaved differently.<sup>5</sup> I propose that we reframe this concept of “incentivized actions” to include “incentivized desires.” I will argue that if an agent can modify her desires and actions based on the different incentives and is fundamentally sane, she is free.

For example, suppose the alcoholic was given a bribe to stop drinking, or was threatened with jail for her actions. The alcoholic, being an irreversible addict, still chooses to drink given an arbitrarily high number of disincentives and thus does not have the control over her desires required for her actions to be considered free. By contrast, the procrastinator is free because she

---

<sup>5</sup> Section Handout 10/30-31.

can alter her desire to stay on task given sufficient incentives. Thus, the concept of alterable desires separates the cases of overwhelming compulsions and moderate desires.

To see why our condition should not require one's actions to be consistent with one's first-order desires, consider the example of a disciplined agent who aspires to eat healthily despite experiencing first-order desires to eat unhealthy foods. If freedom required one's actions to match one's first-order desires, eating healthily against one's temptations would not be considered free, and our moral pride in "self-discipline" would cease to exist. But what about the agent who is physically forced to commit a crime against her will? With our conditions, we ascertain that our agent would not have acted differently from this state of the world, even when given different incentives. Thus, our revised condition allows for an expanded conception of moral responsibility without implicating those who could not have changed their actions.

One could argue that our new looser condition fails to address cases of alternate states of consciousness — after all, a hypnotized person could theoretically modify her desires based on incentives. However, because the hypnotized person does not have an accurate conception of reality in the moment, she fails to be free in the moment on account of sanity. Thus, with our relaxed version of the deep self condition, sanity becomes even more crucial to our interpretation of moral responsibility.

### Section 3: Conclusion

We have thus formed a modified form of compatibilism that responds to The Clash: one whose conception of freedom relies on sanity and incentive-sensitive desires and actions. While I find Wolf's conditions for freedom and moral responsibility largely convincing, they are flawed

in their willingness to accept “lack of freedom” as a valid excuse for everyday failures of willpower. Through the modification that this paper introduces, we begin to tease apart the difference between an agent who truly lacks freedom and one who simply lacks self-control.

WC: 1490

### Bibliography

Wolf, Susan. "Sanity and the Metaphysics of Philosophy." In *Free Will*, edited by Robert Kane.