Value-Laden Free Will

Lixing Miao
University of Pittsburgh

VALUE-LADEN FREE WILL

Lixing Miao
University of Pittsburgh

Abstract

In this paper, I deliver a value-laden concept of free will. This concept of free will depends on an insight into agential psychology: we generally act in light of the desire of the good and the conception of a good life. Since our actions are performed in relation to the motivational final end—a good human life, the value-laden concept of free will defines that an agent has free will in *x*-ing only if the agent is motivated in such a way that *x*-ing is genuinely conducive to a good life. The value-laden concept of free will does not aim at offering a necessary condition of moral responsibility because, intuitively, free will and moral responsibility can come apart in some cases where the wrongdoer is ignorant or deluded. The value-laden concept of free will is able to explain why agents in the following three types of situations are not free: (1) the action is coerced, (2) the action is deluded or ignorant, and (3) there exists some inner hindrance to the agent's reactivity to right reasons. It turns out that the value-laden freedom requires an agent to have a good understanding of the situation in which one acts and the right reasoning concerning what actions count towards a good life.

The value-laden concept of free will depends on a crucial assumption about the motivational structure of agency. I share an insight into human agency purported by Aristotle and rediscovered by modern commentaries (Vogt 2017), that we really want our lives to go well. Generally, we act in light of the desire of the good and the conception of a good life. A genuinely good human life is the motivational final end, in relation to which an agent engages in various actions, from mundane ones, such as getting up regularly every morning, to eventful ones, such as running for office. For the purpose of offering a value-laden concept of free will, it is worth laying out some basic metaethical assumptions I have. For example, I believe that there really are facts about what is good for a given agent. In other words, an agent can be wrong about what actually contributes to a good life, though the right conception of a good life varies from person to person.

Now, what is a value-laden concept of free will? Why should free will be value-laden? Since the desire of a good life is inherent in a motivation, the value-laden concept of free will specifies that an action is free willed only if it is actually conducive to the agent's final end, i.e. a good life. As a result, free will requires the kind of control over one's own actions that actually leads one to live up to the aim at the good.

At first glance, the value-laden concept of free will is surprisingly demanding. Not all participants of the free will debate would agree that free will requires such a high degree of agential control. Many philosophers, for example, have treated free will as the kind of agential control that is necessary for moral responsibility. Those who work with this responsibility-dependent concept of free will may find the value-laden concept of free will unable to explain

a wrongdoer's moral responsibility. If free willed actions, according to my construal, all hit the target of a good life, then no one could perform morally incorrect actions with free will because morally incorrect actions are by no means conducive to a good life. Since the value-laden free will precludes morally incorrect actions, it cannot constitute a necessary condition for the wrongdoer's moral responsibility. It seems that the concept of free will derived from moral responsibility should consist of a moderate degree of agential control that can accommodate incorrect actions. The degree of control that value-laden free will requires is too high to deliver the kind of control that wrongdoer's moral responsibility requires.

I accept that the value-laden free will is not meant to explain wrongdoer's moral responsibility, but I do not think that this is a drawback of my account. As I will show later in the paper, moral responsibility and free will can come apart. In some cases, although the wrongdoers intuitively do not enjoy free will, they can be nevertheless held morally responsible. The dissociation between free will and moral responsibility relies on a crucial feature of agential psychology: our motivations generally aim at what is genuinely good for us, but they may or may not hit the target. This dissociation opens up space for a concept of free will that is independent of moral responsibility. Such a responsibility-independent concept of free will would be value-laden precisely because our actions and motivations are seldom value-neutral.

In order to portray the exact conception of value-laden free will, I will characterize three types of scenario in which an agent fails to hit the target—the final good. Since these agents are motivated in such a way that their actions run afoul of the motivational final good, they can hardly be counted as free. As a result, a value-laden conception of free will should not be too loose, allowing someone in the following three kinds of situations to be counted as a free agent.

(1)   The action is coerced.

The classic example of a coerced action is to be robbed. Imagine a robber points a gun at me and says, "Give me your cellphone!" In this scenario, I would very likely give up my cellphone. If you ask me, "Why did you give your cellphone?" I will answer, "Because I wanted to save my life!" It may seem that my giving the cellphone was a prudent choice when facing such a threat to my life. However, giving up my cellphone in this scenario is hardly a free willed action. If one lives a life where many actions are performed due to systematic coercion, we would not believe that one enjoys free will in these coerced actions because these actions come into conflict with a well-going life. The very fact that an action is coerced is an indication that the agent is wronged by the coercer, which undermines the agent's well-being. In fact, a coercion need not be as dramatic as a threat to life, nor does a coercion need to deprive the agent of the alternatives to do other than what the coercer wishes the agent to do. As Ayer (1954) characterizes, what it takes to be a coercion is that the coercer "should induce me to do what he wants by making it clear to me that, if I do not, he will bring about some situation that I regard as even more undesirable than the consequences of the action that he wishes me to do" (Ayer 1954: 275). In a coerced action, the agent is motivated by the reasonable fear of the threat, and what the coerced agent does is what any rational agent would very likely choose in a similar situation. A coerced agent does what the coercer wishes only to avoid the further undesirable and unjust scenario that the coercer would bring about. If the coercion does not exist, the agent would not perform the undesirable action. For example, if the robber did not present a threat to my life, I would not give up my cellphone but would just run away. The action that the coercer wishes the agent to perform is an intrusion on the agent's course of well-going life because it is not

something the agent would normally choose to do if the coercion did not exist. For this reason, in the *Nicomachean Ethics*, Book III, Chap. I, Aristotle thinks that coerced actions are not chosen "in itself" (1110a). Even though coerced actions are aimed at the final end that is beyond the actions themselves, the coerced agent fails to hit the target of a well-going life. That the agent goes against a good life in the coerced action is, I believe, what underlies the intuition that the agent's free will is undermined.[1]

(2)   The action is deluded or ignorant.

There are two manners in which an agent can be deluded or ignorant in an action: (a) the agent is ignorant of or wrong about some crucial facts about the situation in which they act—facts that the agent would be very interested in because of their relevance to the purpose of the action; (b) the agent is ignorant of or wrong about some ethical implications of their action. These two manners roughly correspond with the two types of agential ignorance that Aristotle defines: the ignorance of the particular and the ignorance of the universal (1110b-1111a). I illustrate these two types of ignorance with the following examples.

Example 1: The sheriff *A* mistakenly arrests an innocent person because of her ignorance of the fact that the fingerprint she takes to be the murderer's actually belongs to the innocent person. After learning that she made a mistake, she feels regretful and very sorry for that wrongly arrested person. I claim that the sheriff *A* can held morally responsible for mistakenly arresting the innocent person if "to be morally responsible" means to be the suitable target of certain interpersonal attitudes such as resentment, gratitude, moral praise and blame (see Strawson 1967/1974). It is very possible that the wrongly arrested person has strong emotional reactions to the sheriff *A*, which is a sign of attributing moral responsibility.[2] Although the sheriff *A* can be held responsible for this wrong arrest, we intuitively believe that she is not free in her decision to arrest the innocent person. This intuition has very little to do with the modality of the sheriff *A*'s action. Rather, it is because the sheriff *A* is constrained by her ignorance of a crucial fact about who the owner of the fingerprint is, a fact that is relevant to her interest in arresting the rightly suspected person. The sheriff *A* is deluded in her action since she is ignorant of the fact about an instrumentally crucial aspect of the particular situation in which she acts. Presumably, the sheriff *A* aims at arresting the actual murderer and fulfilling her job-specific duty, which she values as a crucial component of a well-going life. Unfortunately, her action not only fails to live up to her conception of a good life but also unwittingly undermines her goodness because, in any case, making a great mistake in the job and wronging others really worsen the agent's life.

Example 2: Serial killings happen in a town, but the police are not able to stop the killings because they have not found the murderer. An angry mob gathers in front of the police

---

1   Some might claim that coerced actions can sometimes be conducive to the final good because the agent aims to avoid the more undesirable and unjust outcome that the coercer would inflict. I agree that the aversion to what is more undesirable and unjust is aimed at the final good, nevertheless, I find it very difficult to say that a coerced action can be conducive to a good life because the coerced agent is wronged by the coercer. It seems outrageous to ask the wronged agent to see how the coerced action as such can possibly contribute to a good life. In a coerced action, the agent's desire of the good is co-opted by the coercer. That is, the coercer exploits the agent's desire of the good to induce a state of affairs that undermines the agent's well-going life. Therefore, one who desires a good life would prefer to not confront the coercion in the first place.

2   It is understandable and suitable for the innocent person to feel anger and resentment towards whoever mistakenly arrests her because these emotional reactions are in accord with the ethical fact the she is wronged. I am sympathetic with Strawson's view that "our natural human commitment to ordinary inter-personal attitudes" (Strawson 1962/1974) is such an important component of human life that we cannot and should not get rid of it.

station and threatens that if the police do not arrest the murderer in three days, they will burn down the town hall. The sheriff *B* foresees that a riot would bring about massive casualties if the mob is not satisfied. Therefore, in order to prevent the foreseen riot, he randomly arrests an innocent person to pacify the mob. Presumably, the sheriff *B* wants to protect the civilians of the town by preventing the riot, which he takes as part of his duty, the fulfillment of which he values as a contribution to a good life. Unfortunately, he is deluded in believing that pacifying the mob by arresting an innocent person is the right thing to do. He is ignorant of the ethical fact that he deeply wrongs that innocent person, which immediately not only makes him miss the target of a good life but also erodes his goodness.

Just like the sheriff *A*, the sheriff *B* is unfree because his action runs afoul of his final end—a well-going life. Unlike the sheriff *A*, however, the sheriff *B* is not ignorant of the particular situation in which he acted, but he is ignorant of the universal principle of justice. In both examples, ignorance and delusion in their motivations stand in the way of the agents' pursuits of truly well-going lives, which best explains our intuition that neither the sheriff *A* nor the sheriff *B* enjoys free will in their deluded or ignorant actions.[3]

(3) There exists some inner hindrance to the agent's reactivity to right reasons.

The reactivity to right reasons requires the agent's reliable disposition to act according to right reasons that tell the agent what really contributes to the final good. The failure to manifest this reactivity stems from factors either external or internal to the agent. Since the external constraints of the reactivity to right reasons have to do with the freedom of action rather than the freedom of will, I will focus on the internal hindrance here. Roughly, I think there are two types of inner hindrance to the reactivity to right reasons. (a) *Akrasia* (weakness of will) prevents an agent from doing what is good for one's life as a whole. Someone with *akrasia* has access to the right reasons, but they give in to choices that are immediately more agreeable since they have not built a stable disposition to act for the right reasons. Someone who binge-watches TV shows on Netflix may be well aware that they are procrastinating the important work that they promised to finish on time, but the immediate pleasure of binge-watching overwhelms the right reasoning. For this self-indulgent agent, *akrasia* is an obstacle to live up to his conception of a good life that involves keeping some important promises. We have a strong intuition that akratic agents lack free will because the they fail to set themselves on the courses of activities that are conducive to well-going lives. In other words, the impulsive desire that motivates an akratic action misses the target of the final good. These impulsive desires are not endorsed by the right reasoning, and they stand in the way of the fulfilment of the activities that contribute to the good life. In this way, the inner obstacles, such as impulsive desires, block the fulfillment of the final end that is immanent in the agent's motivational structure. But it is the characteristic of *akrasia* that the agent is able to recover from self-indulgence by adopting the right reasoning through external aids, say, a friend's admonishment. (b) There exists some pathological hindrance to the agent's reactivity to the right reasoning, such as addiction, kleptomania, obsessive-compulsive disorder, etc. What is

---

3   Those who agree that one is unfree when acting through the ignorance of some instrumental facts but insist that one can freely commit wrongdoing due to the ignorance of some ethical facts owe us a story of what justifies the different treatments. One response might be that instrumental facts and ethical facts are different types of facts in terms of the ways in which we get to know them: we learn instrumental facts through empirical observation, but ethical facts cannot be acquired simply by empirical observation. I agree with the distinction between these two types of facts, but this distinction does not support the different treatments. After all, the ignorance of both types of facts are obstacles to hitting the aim inherent in agential psychology—the final good.

peculiar to the pathological hindrance is that the agent experiences very resilient desire of *x*-ing even if *x*-ing is not endorsed by their reasoning of what constitutes a good life. The victims of this type of pathological hindrance cannot retrieve control over their desires as easily as someone who has *akrasia* does.

To rule out the three types of unfree agents mentioned above, I claim that an agent has free will when *x*-ing only if the agent is motivated in such a way that *x*-ing is genuinely conducive to a good life for the agent. That is, free willed actions all hit the target of the final good, which is the aim inherent in an agent's motivational structure. Given the situations that a value-laden conception of free will should rule out, the value-laden free will requires an motivation to accord with (1) a good understanding of the situation in which one acts and (2) the right reasoning concerning what actions count towards a good life.

By "a good understanding" I mean to know some crucial facts and techne that are relevant to the purpose of the action in question. For example, when I freely type on my laptop to finish the final paper of a philosophy course, I must know, among other things, the fact about the topic on the instructor's prompt, the fact about the deadline of the paper, and how to type. Otherwise, my writing the final paper might fail to fulfill the intended purpose—to write on the right topic and to finish it on time, which I consider a component of a well-going life for me. Without the knowledge of some relevant facts and skills, I might write on the wrong topic or fail to turn it in time due to my ignorance of the deadline and my clumsy typing. Many other facts are also involved in writing the final paper, but I do not need to know them when freely writing the paper. For example, I am ignorant of the mechanisms of the keyboard on which I type. I simply have no idea how the circuit inside the keyboard works because I am far from an expert in mechanics and electronics, but my ignorance of these technical issues does not undermine my free will in writing the paper because they are not relevant to the purpose of my action. To know these technical issues might be crucial for free will if they are relevant to the purpose of the action that I value as a component of a good life. For example, suppose I value DIY as something that enriches my life, and my work heavily depends on my laptop. In light of the DIY conception of good life, I try to repair the broken keyboard of my laptop on my own. The cluelessness of some technical issues the lack of certain techne may lead to my failure to repair the keyboard, or what's even worse, circuit damage that ruins my laptop. In this scenario, my ignorance of certain electronical facts and skills undermines my free will in my attempt to repair the keyboard because my action runs foul of the final end of my motivational structure.

By "good reasoning," I mean the reasoning that gets right the conception of a good life and what actions can genuinely count towards a good life.[4] Of course, it is not always easy to deliver a precise conception of a good life because our self-images and self-expectations more or less shift during our lives. The conception of a good life is not something that an agent can fully grasp once for all, nor does it remain the same during one's lifetime. Nevertheless, there really are boundaries of the right conception of a good life. Outside those boundaries, there are actions, such as morally bad ones, that cannot count towards a genuinely good human life. Since the boundaries of a good life are real, an agent can be wrong about the conception of a good life. One who systematically lacks the right reasoning ("an insane agent") fails to achieve

---

4   The right conception of a good life and what actions can genuinely count towards a good life are interdependent. This is because the conceptions of a good life involve small- and mid-scale activities as its substantial contents. A conception of a good life without any implication of what can count towards a good life is arguably empty and ineffective. For the discussion on the interrelation between activities and the final end, see Vogt 2017, 129-30.

a right conception of the good, but the distorted conception of a good life is still in effect in one's motivational structure. An insane agent may still want her life to go well, but her distorted conception of a good life makes her fail to live up to the final good—a good life. According to my construal of free will, someone who lacks the right reasoning cannot enjoy free will in the resultant actions of the distorted conception of a good life, because these resultant actions always miss the target of the motivational final good.

If the value-laden concept of free will makes sense, then whether an action is performed with free will is independent of whether the action is determined or not. Even if the action is causally determined, it is possible for an agent to be motivated in such a way that her actions are genuinely conducive to the motivational final end. Under the value-laden concept of free will, compatibilism is true.

References

Aristotle. (2000). *Nicomachean Ethics*. Trans. R. Crisp. Cambridge: Cambridge University Press.

Ayer, A. J. (1954). Freedom and necessity. *Philosophical Essays* (1954): 271-84.

Strawson, P. F. (1962/1974). Freedom and resentment. In *Freedom and Resentment and Other Essays*, 1-25. Methuen & Co. Ltd.

Vogt, K. M. (2017). *Desiring the Good: Ancient Proposals and Contemporary Theory*. Oxford University Press