

Francisca Vasconcelos

TA: Cosmo Grant

24.00 - Problems of Philosophy

4/26/2018

**Prompt:** *Strawson argues that “true or ultimate moral responsibility is impossible”. What does he mean by that? Does his conclusion have any implications for how we should act, or penalize bad behavior? Set out and critically examine his argument. Has he established his conclusion?*

## No Such Thing as Moral Responsibility

In *Free Will*, Galen Strawson proposes the “Basic Argument,” claiming that, regardless of determinism, “we can never be truly or ultimately morally responsible for our actions.” In this paper, I will summarize and analyze Strawson’s argument. Furthermore, I will refute the claim that his theory is independent of determinism and assess its potential societal implications.

Let us begin by reviewing Strawson’s argument. He starts with an example of choosing whether to buy cake or donate to famine relief. While most people would believe that in this situation they are making a choice based on free will, Strawson disagrees. He instead proposes the “Basic Argument,” which he claims is independent of whether or not the universe is deterministic. He essentially argues that because nothing can be *causa sui* (the cause of itself), no person can be held morally responsible for their actions. His logic goes as follows. At any given time, a decision we make is determined by our current mental state. Thus, in order to be responsible for this decision, we must be responsible for our current mental state. However, our current mental state is a function of our previous decisions, which depended on our previous mental states. In Strawson’s own words, “Here we are setting out on a regress that we cannot stop. True or ultimate self-determination is impossible because it requires the actual completion of an infinite series of choices of principles of choice.” In other words, he claims that we cannot be held responsible for our decisions, because they are ultimately a result of our “heredity and previous experience,” over which we had no control. Strawson then proceeds to discuss ultimate moral responsibility and notes that people generally feel responsible for the way they are. This prompts him to reformulate his “Basic Argument”, as follows:

**Premise 1.** *You do what you do because of the way you are.*

**Premise 2.** *To be truly responsible for the way you are, you must have intentionally brought it about that you are the way you are, and this is impossible.*

**Conclusion 1.** *You can’t be truly responsible for the way you are.*

**Premise 3.** *To be truly morally responsible for what you do, you must be truly responsible for the way you are – at least in certain crucial mental respects.*

**Conclusion 2.** *You can't be truly responsible for what you do.*

With this argument, Strawson claims that because no one is morally responsible, punishment and reward should not be used. He goes so far as to claim, “it is exactly as just to punish or reward people for their actions as it is to punish or reward them for the (natural) color of their hair or the (natural) shape of their faces.” Finally, Strawson refutes counterarguments to his “Basic Argument,” claiming that his theory is not widely adopted because “the conviction that self-conscious awareness of one’s situation can be a sufficient foundation of strong free will is very powerful.”

As a strong determinist, I, like Strawson, do not believe in free will nor ultimate moral responsibility. However, my reasons for believing so are based entirely on the laws of physics. Were we to have a complete understanding of the law of physics and somehow have knowledge of the state and position of all the particles in the universe, I argue that we could calculate the motion of the particles at any future (or past) time, thus allowing us to predict what will happen at any given time. Some argue that because modern quantum physics relies on stochastic processes, there is an element of free will in the belief that we can somehow control the “collapse of the wavefunction” of the particles that make up our mind/body. However, I strongly disagree with this for two reasons. First, while quantum physics has helped us explain a lot of physical phenomenon, I do not believe it will lead to a Grand Unified Theory<sup>1</sup>, because there are still too many holes in our knowledge that cannot be explained by it (i.e. the inability to explain gravity). Just as Newtonian physics was a simplification of Einsteinian physics, I believe that we will develop a deeper theory of quantum mechanics and that these processes which currently seem probabilistic will have an underlying deterministic cause. However, even if the universe turns out to be stochastic, I still would argue against the notion of free will. Given the state of all the particles in the universe, there are a finite number of states (with varying likelihoods) that these particles can be in at a future time. The notion that we could somehow influence the ultimate state of the particle seems absurd. Each quanta in a human body

---

<sup>1</sup>[https://en.wikipedia.org/wiki/Grand\\_Unified\\_Theory](https://en.wikipedia.org/wiki/Grand_Unified_Theory)

has a wavefunction, which defines its probabilistic distribution of states when it collapses. Since the collapse is an entirely stochastic process, the only way we could influence the eventual state of the particle would be by altering the wavefunction of the particle. However, I would argue that the distributions of these wavefunctions (before the collapse) are deterministic and defined entirely by the previous state of the universe. Although for different reasons, Strawson as well agrees that it is “absurd to suppose that indeterministic or random factors, for which one is obviously not responsible, can contribute in any way to one’s being truly or ultimately morally responsible for how one is.”

We discuss this determinist view because there is actually a strong parallel to Strawson’s “Basic Argument.” Let us consider the cake and famine relief example. A determinist argues that the decision made has been determined since the Big Bang, which established the first state of the universe, from which all other states follow via the governing laws of physics. Strawson, on the other hand, argues that our decision to buy cake has been determined since our birth, which was when our first mental state was defined, from which all other mental states resulted via the governing decisions and “outside of our control” factors. However, if we analyze Strawson’s argument more deeply, how was our state at birth determined? It was determined by the state of our mother when we were conceived. Our mother’s state at the time can then be traced back to her state of mind at birth. Continuing this progression, we hone in the state of mind of the first human, then the state of mind of the first monkey and eventually to organisms that had no mind and, thus, no state of mind. However, we can continue this train of logic, since Strawson claims that “when one acts for a reason, what one does is a function of how one is, mentally speaking. (It’s also a function of one’s height, one’s strength, one’s place and time, and so on; but the mental factors are crucial when moral responsibility is in question.)” Thus we can trace our current mental state back to the state of the first eukaryote, which was determined by the chemical composition of the universe that resulted in this first living creature. This was ultimately determined by the state of the universe at the Big Bang. Thus, I strengthen my claim. There is not just a strong parallel between Strawson’s argument

and that of determinism, they are actually one and the same. I would claim that Strawson's "Basic Argument" is just a high-level, simplified, disguised form of the determinist argument. Thus, it is wrong for Strawson to assert that his view is independent of determinism. In fact, from a determinist view, could we not just interpret 'mental states' and 'decisions' as a specific arrangement of atoms in the universe?

Since I am a proponent of determinism, which I argue is the basis of Strawson's theory, I agree with his assertion that there is no moral responsibility. However, he claims that "the conviction that self-conscious awareness of one's situation can be a sufficient foundation of strong free will is very powerful. It runs deeper than rational argument, and it survives untouched, in the everyday conduct of life, even after the validity of the Basic Argument has been admitted." Thus, Strawson believes that even if people were to accept his logic, no one would truly consider it, because the belief in our own free will is too strong. However, I argue that society and the modern legal system already take, at least a weak version of, Strawson's argument into consideration. For example, when someone commits homicide, they are not immediately given the death penalty. Instead, they are taken to court, where the circumstances surrounding the murder are considered. The guilty party can plead lencency on the premise of mental health or other causes outside of their control (many of which have to do with the 'heredity' and 'environmental factors' discussed by Strawson). A clear example of this is the "Afluenza" trial, which took place a few years ago. The defendant killed several people while drunk driving, but came out of the trial innocent, having argued that he was raised to be so spoiled he could not distinguish right from wrong. While some may disagree with the outcome and implications of this trial, it is clear that our legal system takes into account factors outside of our control that may lead us to make certain wrong decisions. On the other end of the spectrum is rewarding people for reasons outside of their control, which is also already implemented in our society. For example, take the system of affirmative action used by most American universities. Given two applicants to the same university with equal talent and resumes, the spot will be given to

the poorer/minority student. This student, thus, is being rewarded for the fact that he is poor or in an underrepresented population, both of which are factors entirely out of the control of the student.

It should also be mentioned that punishment for bad behavior is not uniquely motivated by ideas of moral responsibility. Without it, there would be some risk of society descending into total anarchy, which would not work for most people. Hence, I do not believe that we can take Strawson's argument to the extreme of disbanding all forms of reward and punishment. Independently of whether we believe in moral responsibility, there is a need to penalize behavior that could create chaos. At the end of the day, it does not matter as much if we are morally responsible or not. Punishment is just a set of rules that keeps in prison those most prone to commit certain chaotic acts. If you kill someone you go to jail. This is no different from the laws of nature. If you jump out of the window you will die. There is no "moral responsibility" here, it does not matter if you meant to jump or not. The result is the same. There are some "attenuating circumstances", e.g. if there is a mattress outside, you will only get sore. The same happens for punishment. If society believes that you did not mean to do it, it will soften your landing with the mattress of a lesser penalty. This just guarantees that those most likely to commit the act again will be more severely penalized. In this way, the probability of chaos is minimized. One could argue that moral responsibility is just a construct that society has evolved to enforce this idea. Hence, while there is merit in the discussion of moral responsibility, the conclusions cannot be used as a pretext to simply disband penalties for bad behavior. Some other construct that guarantees the same outcome (minimum probability of chaos) must be put in place.